دار المنظومة
DAR ALMANDUMAH
الرواد في قواعد المعلومات العربية

العنوان:	التحيز في نظم استرجاع المعلومات على الويب : دراسة تطبيقية لتأثير التوجهات السياسية والاقتصادية في محركات البحث
المصدر:	اعلم
الناشر:	الاتحاد العربى للمكتبات والمعلومات
المؤلف الرئيسي:	إبراهيم، سيد ربيع سيد
المجلد/العدد:	ع22
محكمة:	نعم
التاريخ الميلادي:	2018
الشهر:	يونيو
الصفحات:	137 - 196
رقم MD:	970655
نوع المحتوى:	بحوث ومقالات
اللغة:	Arabic
قواعد المعلومات:	HumanIndex
مواضيع:	التحيز المعلوماتى، الويب ونظم الاسترجاع المتحيزة، محركات البحث
رابط:	http://search.mandumah.com/Record/970655

© 2021 دار المنظومة. جميع الحقوق محفوظة.
هذه المادة متاحة بناء على الإتفاق الموقع مع أصحاب حقوق النشر، علما أن جميع حقوق النشر محفوظة. يمكنك تحميل أو طباعة هذه المادة للاستخدام الشخصي فقط، ويمنع النسخ أو التحويل أو النشر عبر أي وسيلة (مثل مواقع الانترنت أو البريد الالكتروني) دون تصريح خطي من أصحاب حقوق النشر أو دار المنظومة.

التحيز في نظم استرجاع المعلومات على الويب
دراسة تطبيقية لتأثير التوجهات السياسية والاقتصادية في محركات البحث

د. سيد ربيع سيد إبراهيم
مدرس نظم استرجاع المعلومات
جامعة بني سويف– كلية الآداب

مستخلص:

تضمنت الويب مواد معلومات نوعية عكست مختلف الأيديولوجيات والرؤى الفكرية والتوجهات المختلفة، حيث قدمت قدرات النشر الرقمي إمكانات التعبير عن التوجهات الأيديولوجية والسياسية والمنافسة الاقتصادية، ومن ثم حملت مواد الويب الموجهة فكريا ضمن محتوياتها خصائص التحيز. وقد تأثرت نظم استرجاع الويب بعوامل التوجه والتحيز النوعي سواء الاقتصادي الدعائي أو السياسي الحزبي أو الأيديولوجي، بما انعكس بتبعيته على جوانب تنظيم وبحث واسترجاع مواد الويب. وقد استهدفت هذه الدراسة البحث في خصائص التحيز كظاهرة طرأت على نظم استرجاع الويب وغيرت من طبيعة أنشطتها وإجراءاتها بما أثر بدوره على طبيعة النتائج المسترجعة الموجهة لمستخدمي هذه النظم. وقد شملت الدراسة تحليل آليات وبرمجيات العمل بنظم الاسترجاع المتحيزة، والتغيرات التي تطرأ عليها فور فرضها نوعا من التحكم أو التدخل في نوعية وترتيب النتائج المسترجعة. وقد استخدمت الدراسة في إجرائها لدراسة تطبيقية عينة من أعلى نظم الاسترجاع استخداما وفقا لتصنيف **alexa**. وخرجت الدراسة بنتائج من أهمها أن نظم استرجاع الويب لم تعد نظم المعالجة والبحث لمواد الويب بشكل محايد من منظور دراسات الإفادة في مجال المعلومات، بل ينظر الآن لنظم الاسترجاع باعتبارها نظم إدارة وتداول المعرفة الرقمية بما يوافق سياسات عمل ممنهجة. فلا يقتصر الأمر الآن في نظام الاسترجاع على اتاحة مادة الويب وفقا لما هي عليه، إنما يتعدى ذلك إلى تخير موضع وموضوع الإتاحة وصب المحتوى في معين محدد.

الكلمات المفتاحية: الويب المتحيزة، تحيز نظم الاسترجاع، التحيز، البحث الاسترجاع بالتحيز.

المقدمة المنهجية

1/0 تمهيد

استخدم الإنسان مصادر المعلومات كرسالة فكرية تحمل معلومات حول المعتقدات والحقائق وكل ما من شأنه إحداث تغيير في الحالة المعرفية والوجدانية لمتلقي الرسالة. وقد اهتمت علوم المكتبات والمعلومات بدراسة توجهات واتجاهات الرسالة الفكرية المتضمنة في القالب المادي أو الشكلي لمصدر المعلومات، لذا فكانت عمليات تحليل المحتوى والتنظيم على أساس الموضوع والمجال الفكري، ثم عمليات الرقابة والتحليل للمضمون الفكري. وعملت مؤسسات المعلومات منذ بدايتها على التحري في عمليات الانتقاء والاختيار لمصادر المعلومات

اعتمادا على طبيعة التوجه المجتمعي لمستخدميها، فضلا على تقييد عمليات الاختيار والاضافة بالمرجعيات الثقافية والعقائدية التي تنتمي لها مؤسسات المعلومات. وتشير كل هذه الحيثيات إلى تأصيل فكرة حمل مصدر المعلومات لمضمون فكري موجه ومحدد يستهدف معه التأثير في مستخدمه سواء أكان ذلك التأثير إيجابيا أو سلبيا. ولم تترك الايديولجيات العالمية والحركات الفكرية فرصة مخاطبة مريديها من خلال مواد المعلومات التي تنتجها وتبثها هذه الأيديولوجيات، بل وحرص هذه الاتجاهات الفكرية على نشر أفكارها واستقطاب الآخرين من حلال مصادر المعلومات. وكان ذلك لتميز مصادر المعلومات بقدرات الانتشار والانتقال والتوسع حتى في بيئتها التقليدية، فضلا عما أوجدته البيئة الرقمية من قدرات النشر.

قابل النشر الفكري والأيدلوجي لمصادر المعلومات في البيئة التقليدية قدرات أكبر من التحكم والضبط لانتشار واضافة مصادر المعلومات في مؤسسات المعلومات ؛ حيث حضت قوانين وسياسات النشر الوطنية للدول على تحديد قواعد ومعايير اختيار ونشر مصادر المعلومات داخل الدولة. وتغيرت العلاقة بين قدرات النشر وقدرات التحكم باستخدام الويب كبيئة رقمية لنشر مواد المعلومات ؛ حيث تحولت عمليات النشر من الشكل المؤسسي إلى الشكل الشخصي الفردي، فضلا عن قدرات الانتقال والانتشار السريع لمواد معلومات الويب، ناهيك عن امكانات استخدام وسائط الصوت والصورة في حمل المعلومات للنشر السريع لمواد المعلومات التي لم تعد لها حدود مكانية أو زمنية تستطيع معها القدرات القابية السيطرة عليها. وقد تميزت مواد المعلومات في بيئة الويب بإحداث التأثير المباشر المستهدف في نطاق أوسع من المتلقيين للرسالة الفكرية ورواد الفضاء الإلكتروني، فلم يعد نشر الأفكار مرتبطا بعدد مستطيعي القراءة والكتابة أو الاسوياء من مستخدمي مصادر المعلومات، بل تغير شكل قالب الرسالة الفكرية لتصبح في متناول الانسان محدود القدرات التعليمية والحسية من ذوي الاحتياجات الخاصة. ويكمن التأثير الحقيقي لمواد المعلومات الرقمية في وسيط التنظيم والمعالجة والبحث لهذه المواد المتمثل في نظم استرجاع المعلومات، فهي تمثل الرابط بين مستخدمي الويب من ناحية وبين الوصول والعثور على مواد المعلومات من ناحية أخرى.

وتعمل نظم استرجاع مواد الويب على جمع وتنظيم وبث هذه المواد التي تحمل الرسائل الفكرية اعتمادا على امكانات التكشيف واختيار المفردات والواصفات الدالة على محتوى هذه المواد أو ممثلات الرسالة الفكرية المتضمنة في مواد المعلومات. وبناءا على ذلك فقد أضحت نظم استرجاع مواد الويب المتحكم الأكبر في دعم نشر رسائل فكرية وتوجهات عقائدية محددة من خلال إعادة تشكيل الوصف الموضوعي والفكري لمواد المعلومات، بل وتوجيه مستخدميها من خلال توفير آليات بحث واسترجاع موجهة تسعى بالمستفيد إلى استرجاع مواد معلومات محددة، أو أن تكون أداة متحيزة لإيصال مواد معلومات بذاتها إلى قطاع أو نطاق محدد من مستخدميها. وقد تكاملت مع نظم استرجاع المعلومات شبكات التواصل الاجتماعي التي تخاطب نطاقا واسعا من مستخدمي مواد الانترنت والتأثر بمحتواها، غير أن مواقع الشبكات الاجتماعية تختلف عن نظم استرجاع المعلومات في خضوعها

لجوانب التوجيه والتحيز أكثر من نظم استرجاع المعلومات التي تمثل نظما معيارية وقياسية، تخضع على أساسها لضوابط في عمليات المعالجة والتنظيم ويغلب عليها الجانب المعلوماتي في إدارة مواد المعلومات، بما يجعلها أدوات أقرب للحيادية والموضوعية في استرجاع وبحث مواد المعلومات. ومن ثم فإن ثمة معايير وقواعد لأنشطة وإجراءات معالجة مواد المعلومات واسترجاعها يجب أن تتوافر في نظم استرجاع المعلومات، حتى يتسم نظام الاسترجاع بالحيادية والموضوعية وعدم التحيز الفكري أو الإيديولوجي في وصول المستفيد أو تلقيه لمواد معلومات بذاتها لحمله على استقبال أو تلقي فكرة محددة.

2 / 0 مشكلة الدراسة:

عملت الويب كبيئة نشر رقمية على منح فرصة النشر الفردي والجماعي والمؤسسي لكل مستخدميها دون أن تضع هذه البيئة أية قواعد أو ضوابط لعمليات نشر وتدويل مواد المعلومات بوسائطها النوعية. وقد قابل الامكانات اللامحدودة لعمليات النشر قدرات غير محدودة لنظم استرجاع المعلومات في دعم اتاحة ووصول مستخدمي الويب لمواد المعلومات المنشورة في فضائها، غير أن أدوات بحث الويب التي تمثل نظم الاسترجاع لهذه المواد والمعنية بمعالجة وتنظيم وبحث مواد المعلومات تخضع لتوجهات وانحيازات الشركات ورؤوس الأموال المالكة للنظم أو الداعمة لها، وتهتم الدراسة على وجه الخصوص بالتأثيرات والتوجهات السياسية التي يتبناها نظام استرجاع المعلومات، وذلك لما يجري في العالم الآن من تغييرات وصراعات سياسية ذات أبعاد اقتصادية قد تتشابك من قريب أو بعيد مع شركات ورؤوس الموال نظم الاسترجاع. ومن ثم فقد يشوب عمل نظم استرجاع المعلومات تحيزا لفكر معين أو حيادا عن منهج الموضوعية في التعامل مع مواد المعلومات وإجراءات معالجتها واسترجاعها وتقديمها للمستفيدين، وهذا الأمر قد يسره وساعد عليه القدرات التقنية لنظم الاسترجاع في وصف وبحث مواد المعلومات النوعية على الويب. وبناء على ذلك فإن نظم الاسترجاع تحتاج إلى معايير وقواعد قياس الأداء التي تحدد توجهاتها في معالجة وبث المعلومات، وتكمن مشكلة الدراسة في كيفية ايجاد آليات ومعايير وقواعد لقياس الحيادية والتحيز في نظم استرجاع المعلومات ومن ثم الحكم على أداء نظام استرجاع المعلومات ودقته في المعالجة وعدم الإخلال بجوانب التنظيم والبحث. تتبنى الدراسة فرضية هي (**تحمل نظم استرجاع الويب ملامح تجارية وأيديولوجية وسياسية تمثل توجهات وتحيز للنظام تنعكس بدرجات متفاوتة على نوعية نتائج البحث ومواد الويب المسترجعة لمستخدميها**)

3 / 0 تساؤلات الدراسة :

تتحدد مشكلة الدراسة في التساؤلات التالية :-

1. ما مظاهر التحيز وعدم الحيادية في بحث المعلومات على نظم استرجاع مواد الويب ؟
2. ما إجراءات وآليات قياس الحيادية والموضوعية في أنشطة وعمل نظم استرجاع المعلومات ؟

3. ما مدى تأثر نظم استرجاع الويب بالتوجهات السياسية والأيديولوجية للشركات المالكة لها ؟
4. ما المعايير التي تقنن وتضبط الموضوعية والحيادية في استرجاع المعلومات على الويب ؟
5. كيف يمكن استخدام مؤشرات بحث واسترجاع المعلومات كمقياس لحيادية أو تحيز محركات البحث ؟

0 / 4 اهداف الدراسة :

تسعى الدراسة لتحقيق الأهداف التالية :-

1) دراسة مظاهر التحيز وعدم الحيادية في بحث المعلومات على نظم استرجاع مواد الويب ؟
2) تحديد إجراءات وآليات قياس الحيادية والموضوعية في أنشطة وعمل نظم استرجاع المعلومات ؟
3) قياس مدى تأثر نظم استرجاع الويب بالتوجهات السياسية والأيديولوجية للشركات المالكة لها ؟
4) وضع تصور لمعايير تقنين وضبط الموضوعية والحيادية في استرجاع المعلومات على الويب ؟
5) طرح مؤشرات محددة في بحث واسترجاع المعلومات كمقياس لحيادية أو تحيز محركات البحث ؟

0 / 5 أهمية الدراسة :

تكتسب الدراسة أهميتها من الجوانب التالية :-

1. تمثل الويب بيئة نشر مفتوحة سواء من قبل منتجي مواد المعلومات ومرسلها أو من قبل متلقي المعلومات والباحثين عنها، لذا فقد وفرت بيئة الويب امكانات نشر الأفكار والتوجهات لمدى ونطاق أوسع. وهو ما يجعل أدوات بحث الويب أمام فيضا هائلا من التيارات الفكرية والايديولوجية وعليها التحكم بها حتى الوصول الى مستخدميها.
2. تمثل نظم الاسترجاع حلقة الوصل بين مواد المعلومات النوعية على الويب من ناحية وبين مستخدمي الويب من ناحية أخرى، ما يجعلها أمام تحدٍ كبير في الحفاظ على حيادية المعالجة والتنظيم لمواد معلومات نوعية هي في ذاتها غير محايدة ومتحيزة لتيار محدد.
3. تستخدم نظم استرجاع المعلومات آليات ومعالجات لتنظيم وبحث مواد المعلومات الرقمية تستطيع بها تقديم وتأخير مرتبة كل مادة وتغيير الوزن النسبي لأهمية وعلاقة مادة المعلومات بموضوع البحث، بما يجعل المواد المسترجعة مرتبة بشكل محدد امام المستخدم. وإذا لم تحافظ نظم الاسترجاع على مبدأ الحيادية في التنظيم والبحث فستمثل هذه الأدوات أبواقا لتيارات وأفكار موجهة، في حين أنها بالاصل أدوات لتنظيم ومعالجة المعلومات للمحتوى على إطلاقه دون تمييز لموضوع أو فكر على غيره.

4. تبنى نظم استرجاع الويب برأس مال تجاري لشركات مالكة او داعمة لواجهة وبرمجيات عمل النظام، وهو ما يشكك في قدرة واجة العمل المعلوماتية ونظام الاسترجاع على الفصل بين توجهات وتحيز الشركات المالكية وبين أسس الحيادية والموضوعية في تنظيم وبحث مواد المعلومات، واسترجاعها وفقا لمحتواها وآليات تنظيمه، لا وفقا لتمييز او تحيز يوافق توجهات رؤوس الاموال الأيديولوجية أو السياسية.

0 / 6 حدود الدراسة:

- **الحدود النوعية :** تختص الدراسة بالتعامل مع نظم استرجاع المعلومات التي تعمل بشكل آلي خالص متمثلة في محركات بحث الويب.
- **الحدود الموضوعية :** تركز الدراسة على طبيعة التحيز كهوية اداء داخل نظم الاسترجاع والجوانب الداعمة والمنعكسة على نتائج البحث.
- **الحدود الزمنية :** تعمل الدراسة على تتبع ملامح التحيز في نظم استرجاع الويب في العقد الأخير واما الدراسة التجريبية فقد أجريت في فترة نوفمبر 2016 إلى مارس 2017.

0 / 7 مصطلحات الدراسة:

- **التحيز:** تحدد الدراسة مفهوم التحيز على أنه تدخل متعمد من جانب نظام استرجاع المعلومات في نوعية وتوزيع وترتيب النتائج المسترجعة بهدف تقديم نتائج على أخرى لصالح توجه محدد يتبناه النظام.
- **نظم استرجاع الويب :** يمثل نظام الاسترجاع كيان رقمي يعمل في بيئة الويب لمعالجة وبحث مواد الويب واتاحتها وفقا لطبيعة ونوعية الحاجات والمجالات الموضوعية والتوجهات الفكرية مسار بحث المستخدم.

0 / 8 منهجية الدراسة:

تتبنى الدراسة استخدام منهج البحث **الوصفي التحليلي** وذلك لدراسة وتحليل ظاهرة التحيز في مواد معلومات الويب كبيئة نشر الكترونية ومدى تأثر نظم الاسترجاع بهذه الظاهرة وقدرتها على ضبط عملها، واستقراء الانتاج الفكري والدراسات المرتبطة بهذه الظاهرة. بالاضافة الى استخدام منهج البحث **التجريبي** وذلك لملاحظة وقياس مدى تأثر نظم استرجاع المعلومات بعوامل التحيز والموضوعية في بحث واسترجاع المعلومات.أما **أدوات البحث** فتتمثل في أداة البحث الوثائقي في جمع الانتاج الفكري المرتبط بظاهرة البحث، واستخدام أداة الابحار والمعايشة مع الويب الى جانب استخدام أداة الملاحظة في الجانب التطبيقي للبحث.

عينة الدراسة : عمل الباحث على تحديد عينة الدراسة وفقا لآخر ترتيب وتحديد من المؤسسات المختصة بالتقييم والترتيب السنوي لمحركات البحث. وقد اختار الباحث ستة محركات تصدرت الترتيب في موقع **alexa**.

0 / 9 الدراسات السابقة:

استخدم الباحث قواعد البيانات العالمية المتضمنة بيانات اتحاد المكتبات الجامعية المصرية **EULC** مثل **springer, science direct, IEEE, Dissertation proquest** وغيرها، حتى يمكن البحث عن الدراسات المثيلة والسابقة التي عنيت بدراسة ظاهرة التحيز وتحليل أداء نظم استرجاع المعلومات على الويب. وقد نتج عن هذا البحث العديد من الدراسات التي تناولت موضوع التحيز والتوجهات السياسية والاقتصادية والايديولوجية التي قد تتبناها محركات البحث على الويب باعتبارها نظم استرجاع المواد الرقمية، مع تحليل سبل التعامل مع هذه الجوانب إذا ما اتصفت بها محركات بحث الويب. أما الانتاج الفكري العرب فقد خلا من الدراسات التي تتناول موضوع التوجهات والتحيز لنظم استرجاع. وقد كانت أهم الدراسات استوقفت الباحث ما يلي :

الدراسة الأولى :

Owens, Ciaran. A Study of the Relative Bias of Web Search Engines Toward News Media Providers, University of Glasgow, 2009, (Master)

ركزت هذه الدراسة على معالجة واسترجاع المواد الإعلامية داخل محركات البحث وكيفية أداء عمل محرك البحث وفقا لتوجه او تحيز يؤثر في طبيعة النتائج المسترجعة بما يؤثر على توجهات مستخدمي الويب، وهو ما يعرف بالبحث في مواد الزمن الحقيقي المرتبطة بالتغيير الوقتي والتحديث المتسارع في مواقع الأخبار. وقد اختارت الدراسة ثلاثة محركات هم **google, bing, yahoo** لدراسة طبيعة المعالجة والنتائج المسترجعة المتأثرة بظاهرة التحيز في محركات البحث الثلاث. وتختلف الدراسة الحالية عن الدراسة السابقة في كونها تشمل مختلف مواد الويب النوعية والتركيز على آليات وبرمجيات العمل في نظم الاسترجاع ومظاهر التحيز، فضلا عن دراسة خصائص النتائج المسترجعة نوعيا وعدديا وفقا لقضايا معاصرة تمثل توجهات وسياسات عمل ممنهجة لنظم الاسترجاع.

الدراسة الثانية :

Diaz , Alejandro M. Through the Google Goggles : Sociopolitical Bias in Search Engine Design, Stanford University, 2005, (PHD)

اهتمت هذه الدراسة بتحليل خصائص العمل في محرك البحث **google** كدراسة حالة لخصائص البحث والاسترجاع على هذا المحرك ؛ حيث وصفت هذه الدراسة نظام **google** بحارس الويب نظرا لما يشغله من مكانة كبيرة في عدد عمليات البحث والاسترجاعى التي تجربعلى مواد الويب. وقد اهتمت هذه الدراسة بطبيعة الخصائص السياسية والاجتماعية التي يعمل بها محرك البحث، التي بدورها تؤثر على طبيعة الآليات والبرمجيات التي يعمل بها **google**. وتختلف الدراسة الحالية عن هذه الدراسة في تتبع ظاهرة التحيز في كل المراحل وإجراءات العمل التي تمر بها مادة المعلومات داخل نظم استرجاع الويب، وانعكاس ذلك على طبيعة وبرمجيات العمل المتحيزة.

0 / 10 **خطوات الدراسة:**

1) دراسة أدبيات الإنتاج الفكري وتحليلها فيما يخص مفاهيم التحيز ومظاهره في نظم الاسترجاع.
2) دراسة واستقراء آليات وإجراءات التحيز في نظم الاسترجاع حال الارتكان إلى جانب نوعي محدد من المعلومات.
3) القياس التطبيقي لآليات ومعايير التحيز او المصداقية التي تستخدمها نظم استرجاع المعلومات في إجراءات البحث.
4) توصيف وتحديد معايير ضبط التحيز خاصة السياسي وقياس تواجده وتأثيره على بحث المعلومات في نظم استرجاع المعلومات.

المبحث الأول: التحيز والموضوعية لمواد الويب ونظم استرجاعها:

تحمل مواد المعلومات الرقمية في محتواها الفكرة والرسالة المعلوماتية الموجهة من مرسلها أو ناشريها في البيئة الرقمية إلى مستقبليها من مستخدمي الويب، وفي هذا المحتوى تكمن التوجهات العلمية أو الفكرية للرسالة أو مادة المعلومات لتقع بتأثيرها ومضمونها على قلب وعقل مستخدم معلومات الويب. وقد تأثرت أدوات تنظيم وبحث معلومات الويب بنفس هذه الفكرة في كونها كيانات ونظم معلوماتية تحمل الأفكار المنشورة في شكل وسائط متعددة إلى الباحثين عنها ومستخدميها ؛ حيث يعمل نظام استرجاع المعلومات على تقديم معلومات محددة لاستفسارات بحثية محددة وهو ما قد يعني تحقق مستوى الدقة المطلوب أو انعدامه في المعلومات المسترجعة الموجهة إلى مستخدمي النظام. مما يعني ذلك وجود فكرة التحيز أو عدم الحيادية في مواد المعلومات وفي نظم الاسترجاع التي تقوم بدورها على إتاحة هذه المواد لمستخدميها. ويتطرق هذا المبحث لمناقشة مفاهيم التحيز في مواد المعلومات وأشكال هذا التحيز وتوجهاته النوعية.

1 / 1 ظاهرة التحيز المعلوماتي : التعريف والأسس:

عنيت كيانات المعلومات الرقمية ومن قبلها التقليدية بدراسة مواد المعلومات من حيث الرسالة الفكرية المتضمنة في هذه المواد والمصادر الموجهة إلى جمهور الباحثين أو المستفيدين من مواد المعلومات. وقد كان الاهتمام بذلك كبيرا في البيئة التقليدية لمصادر المعلومات وازداد لأعلى حالاته في البيئة الرقمية التي أتاحت قدرات النشر الرقمي للأفراد والهيئات. وتعد صفة الحيادية والموضوعية والاهتمام بالحقائق والمعلومات على حساب العقيدة والتوجه هي أصل بنية محتوى مادة المعلومات، غير أن الاختلافات والتنوع الفكري والايديولوجي الذي يشهده العالم فرض نشر مواد معلومات ذات أعداد مليونية تستهدف نشر فكر محدد أو الترويج لمذهب فكري أو عقائدي بذاته. وهذا الفعل يخلق صفة جديدة في مواد المعلومات تسمى التحيز أو عدم الحيادية في مواد المعلومات، وإن كان ذلك عملا مستحبا من وجهة نظر ناشري المعلومات الرقمية، غير أنه يطعن في الثقة والدقة التي تكون عليها مواد المعلومات الرقمية على الويب. ولم يقتصر أمر التحيز هنا في نشر مواد معلومات الويب، وإنما شمل مفهوم التحيز وعدم الحيادية في بيئة الويب أيضا أدوات تنظيم وبحث المعلومات الرقمية على

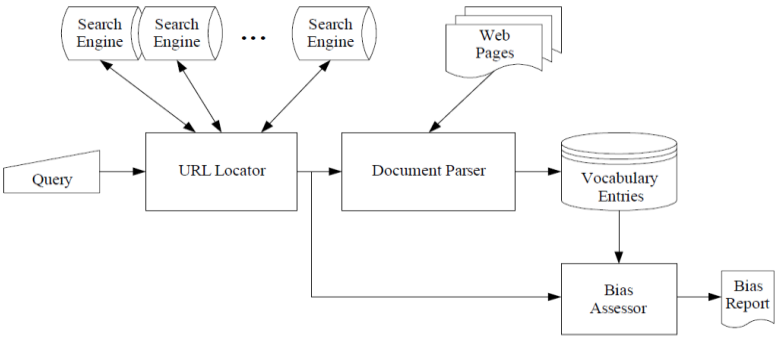
الويب التي تتمثل أفضل صورها في محركات بحث الويب. ويمكن لصفة التحيز أن تطرأ على نظم استرجاع الويب إذا ما عمدت إلى الانتقاء في تجميع مواد معلومات محددة دون غيرها، أو تفضيل مواد معلومات على أخرى في المعالجة والكلمات الواصفة المتحيزة لفكرة محددة، أو كونها عملت على تقديم وتأخير مواد معلومات على غيرها في قائمة النتائج المسترجعة المعروضة أمام المستخدم.

وقد عمدت الدراسة إلى اختيار مصطح التحيز دون غيره كمفردة لغوية للدلالة على عدم الحيادية في نظم استرجاع المعلومات من حيث اختيار ومعالجة وبحث مواد المعلومات ؛ وذلك لأمرين ألولهما : أن مصطلح التحيز هو المقابل اللغوي لمصطلح **Bias** الذي اعتمده الانتاج الفكري الأجنبي للدلالة على ظاهرة التحيز في محركات البحث، ثانيهما : أن مصطلح التحيز لغويا يمثل تعمد التدخل في الاجراء أو الفعل لتغليب أو ترجيح امر على آخر، وهو ما ينعكس نهاية على نتائج محركات ونظم استرجاع معلومات الويب لاسترجاع وتمييز مواد معلومات على أخرى في قائمة المواد المسترجعة وإعطائها ترتيبا أعلى في قائمة المود المستدعاة لاجابة الاستفسارات البحثية. ويختلف ممصطلح التحيز عن الانحياز في كونه إجرائيا ولا يقف عند الانحياز الفكري أو الايدلوجي لتيار أو فكر محدد دون إحداث أثر أو تغيير في مسار عملية الاسترجاع والترتيب لمواد المعلومات في النظام. وقد عرف معجم المعاني والمعجم الوسيط مصطلح التَحَيُّز كمصدر للفعل تَحَيَّزَ ومعناه مُحاباةُ شخص أو مناصرةُ رَأْي على حِساب آخر من دون مُبالاة بالحَقّ والعدالة[1]. وينسحب هذا المفهوم في هذه الدراسة على نظم استرجاع الويب التي تنصر موادا أو محتوى ذات توجه ما على غيره من مواد المعلومات الأخرى المتوافقة وكلمات البحث. أما من وجهة نظم الاسترجاع فالتحيز هو تعمد إدخال مواد معلومات محددة تحمل ترويجا سياسيا أو اقتصاديا للتأثير على رأي المستخدم واقناعه بقبول وجهة نظر محددة[2]. وقد أضاف **Joshua wright** أن ظاهرة التحيز في نظم البحث تتمثل في الفكرة العامة للاختلاف بين النتائج الأصيلة المرتبطة وموضوع البحث بشكل معياري وبين تلك التي تمتلك مزايا مختلفة عن ارتباطها بطبيعة البحث[3]. ويرتبط بمصطلح التحيز مصطلح مقابل هو المصداقية أو الحيادية وهو المصطلح المعبر عن موضوعية نظم استرجاع المعلومات في التعامل مع مواد المعلومات انتقاء واسترجاعا ؛ بحيث يتم اضافة مجموعات المعلومات لقاعدة البيانات دون تحديد مسبق أو تحيز لمسمى موضوعي محدد، أيضا الاسترجاع اعتمادا طبيعة الكلمات الكشفية والدلالية المشتقة من طبيعة النصوص، ثم تركها لمدى صلتها بالموضوع أثناء عمليات البحث والاسترجاع.

وقد تسبب الخضم الهائل من مواد المعلومات الرقمية في وجود التحيز داخل دائرة عمل نظم استرجاع الويب على شكلين مختلفين هما التحيز المتعمد وغير المتعمد **intentional or unintentional** ؛ حيث يتمثل التحيز المتعمد في مجموعات الاعلانات التجارية والمواقع مدفوعة الرسوم حتى توضع في ترتيب متقدم ضمن فائمة النتائج المسترجعة في عمليات البحث داخل نظام الاسترجاع، وهذه تمثل نوعا من التحيز التجاري الربحي الذي يمثل واقع شركات نظم استرجاع المعلومات الرقمية. أما التحيز غير المتعمد فهو يأتي نتيجة للأعداد المليارية من مواد المعلومات التي يتعامل معها النظام معالجة وتنظميا بما يجنح بالكلمات الدالة والمصطلحات الواصفة لمحتوى مواد

المعلومات بإضفاء نوعا من التميز أو احتلال ترتيبا متقدما في مواد المعلومات المتناولة لموضوع ما. ويمثل جانب التحيز صعوبة بالغة على المستخدم في امكانية الملاحظة او التحديد السريع لتوجهات نظام الاسترجاع وعدم حياديته حيال مواد معلومات محددة، وذلك يرجع بالدرجة الألى لكون التحيز كيفي وليس كمي منحصر في مواد بذاتها أو بشكل مميز لعدد معلوم من النتائج المسترجعة. بل إن التحيز من الجوانب التي يصعب تقديرها كميا أمام المستفيد بما يمثل صعوبة أكبر في تحديد درجة التحيز لكل نظام استرجاع على الويب، وهو إذا ما حدث يمكن للمستخدم اختيار أقل نظم الاسترجاع على الويب تحيزا ليستخدمه واثقا فيما يتم استرجاعه. ومن ثم فإن قياس معامل التحيز في نظم الاسترجاع يضع تصنيفا لهذه النظم بحسب الثقة في النتائج المسترجعة ومواد المعلومات المستدعاة حيال موضوع أو مجال ما[4].

إن ثمة مصادر محددة يمكن من خلالها إصباغ محركات البحث بصبغة التحيز مثل ؛ أن تسعى محركات البحث إلى الانتقاء في عمليات الإضافة وتجهيز برامج الزواحف لديها ببعض التعليمات البرمجية بالتركيز على مواد معلومات دون غيرها. وقد يتمثل هذا التركيز في تغطية مكانية محددة أو انتقاء موضوعي أو إضافة المواقع التي لها امتداد محدد او البحث عن مصطلحات بذاتها في عنوان مادة المعلومات **URL**. ومن مصادر التحيز عملية التكشيف ذاتها داخل النظام حيث يبحث النظام عمدا عن كلمات مفتاحية محددة لجعلها كلمات دالة وواصفات كشفية يبنى عليها استرجاع مادة المعلومات، وتتأثر عملية التكشيف بمجموعات الكلمات التي ترد ضمن محارف صفحة الويب أو التي وردت في سياق ومتن المادة ذاتها. كما أن طرح خيارات البحث وتقديم مصطلحات مقترحة لعملية البحث يمثل تحيزا من جانب نظام الاسترجاع في توجيه المستخدم لاختيار كلمات بذاتها ينعكس عليها طبيعة المواد المسترجعة نهاية. ويتبقى الفرز والترتيب كآخر مصادر التحيز في الاسترجاع بحيث يلجأ النظام إلى خوارزميات وآليات تمنح موادا دون أخرى فرصة الترتيب الأعلى في النتائج المسترجعة أمام المستخدم[5]. ويمكن للتحيز أن يأخذ بعدا آخر متمثلا في الاهتمام بأسماء نطاقات محددة **domain name** وهو ما يمثل رمزا لمؤسسات بعينها أو لهيئات ذات توجه محدد، واثبات هذا النوع من التحيز يحتاج إلى اختبار مجموعات النتائج المسترجعة من النظام ورصد نسبة عدد النطاقات المسترجعة **retrieval domains** إلى أعداد النتائج والمواقع المسترجعة جميعها من النظام[6]. وعلى ذلك فإن ثمة نوعين من التحيز يتعلق أحدهما بإدخال وحدات لا علاقة لها باستفسار البحث وتتعلق بأسماء نطاقات محددة، يعلوها التوجه الاقتصادي في المواد المعلوماتية مدفوعة الرسوم أو الايديولوجي لمجموعات المواد التي تحمل رسالة فكرية ما يتبناها محرك البحث. وأما النوع الثاني فهو التحيز من حيث المحتوى المبني على توجيه وتغيير مسار المحتوى من التعبير عن فكرة ما إلى خدمة توجه آخر من خلال تغيير سياسة التكشيف في نظام الاسترجاع[7]. ويعبر شكل (1) عن قياس التحيز في نظم الاسترجاع من منظوري اللإضافة بالانتقاء والتكشيف بالتدخل غير الموضوعي.



شكل (1) قياس التحيز في نظم الاسترجاع من منظوري الإضافة والتكشيف[8]

يصنع التحيز في نظم الاسترجاع أثرا سلبياً على جوانب مختلفة تتعلق بطبيعة مواد ومواقع المعلومات التي تفقد قيمة الرجوع إليها والاستشهاد بمحتواها بما يجعلها تتراجع في تصنيفات مختلفة نتيجة لانخفاض مؤشرات استخدامها. ويأتي ذلك من أن آلية الترتيب في نظام الاسترجاع -**popularity based ranking algorithms** تعمد غالبا إلى تصنيف وترتيب المواقع شائعة الاستخدام إلى فئة أولى تتصدر قائمة النتائج المسترجعة على حساب مواقع ومواد المعلومات الأخرى الأقل شيوعا. ويشمل الأثر السلبي للتحيز مستخدمي النظام أنفسهم حيث يفقدون الوصول إلى النتائج المرجوة الأكثر اتصالا استفسار البحث وموضوعه، بما يدفعهم إلى إعادة البحث أو البعد عن وجهة البحث الصحيحة في استرجاع مواد المعلومات، التي بطبيعتها جاءت متأخرة وغير مرئية لهم بتأخر ترتيبها في قائمة الاسترجاع. وعلى ذلك فنتائج التحيز تنعكس آثارها على المحتوى والرسالة الفكرية لمواد المعلومات لصالح التصنيف الفئوي لمواقع الويب، وعلى دعم النظام المعلوماتي البحثي لنظم الاسترجاع لصالح سياسات الانتماء والربح العائد على الشركات المالكة للنظم[9]. ولذلك تكتسب الحيادية والموضوعية في محركات البحث أهميتها وتأثيرها الإيجابي على المستخدمين بالوصول إلى المعلومات المتصلة بالاستفسار البحثي دون مدخلات إضافية متعمدة من النظام.

2 / 1 الويب ونظم الاسترجاع المتحيزة :

بدأت الويب نشأتها قاصرة على انشاء المعلومات وتقديمها من جانب واحد يمثل أغلبه قطاع المؤسسات والهيئات التي تمارس أنشطة رسمية أو خدمية وفقا لهويتها وطبيعتها. وقد أتاحت الأجيال اللاحقة من الويب قدرات النشر الرقمي للأفراد والمؤسسات المختلفة من خلال منصات نشر مجانية تابعة لبوابات الويب أو مواقع التواصل الاجتماعي، بما سمح لنشر نوعي متعدد الأغراض أتاح لترويج ودعم التوجهات الفكرية والترويج للأفكار والسلع على حد سواء. وقد بلغ عدد المواقع المضافة لفضاء الويب مليار ومائة مليون وثلاثمائة ألف موقع وقت اعداد هذه الدراسة وساعة الكشف عن هذا العدد، وهذا العدد متنامي بشكل متسارع يصل إلى **168,192,000** موقعا سنويا[10]. ومن إمكانات

النشر الواسعة ظهرت مواد معلومات الويب التي تحمل توجها فكريا أو ماديا أو سياسيا قاصدة استقطاب قطاعات مستخدمي الويب للتأثر بمحتواها الفكري ونشره على نطاق أوسع، ومن ثم فإن ذلك قد صنع ما يسمى بالويب المتحيز **biased web** التي تجسدت في مواد ومواقع المعلومات النوعية الإعلامية والشخصية والسياسية والتجارية. ودفع هذا التنامي المضطرد كما ونوعا الهيئات البحثية مثل جامعة ميريلاند **University of Maryland** إلى النظر في مجموعة معايير تحول دون سيطرة المواقع المتحيزة على محتوى الويب، وتمثلت هذه المعايير في أسئلة محددة لقياس درجة التحيز في مادة الويب ومواقعها، مثل[11]:-

- هل المحتوى يستهدف ظاهرة آنية، ومدى درجة الحداثة للرسالة المُضمّنة ؟
- هل يمكن القبول بطلبات تعديل المحتوى أو تحديثه ؟
- من المسئول عن بنية ونشر محتوى ومواد موقع الويب ؟
- ما هي الاغراض التي يستهدفها الموقع من وصوله إلى مستخدمي الويب ؟
- ما هي فئات الجمهور التي يستهدفها هذا الموقع ؟
- ما هي نوعية المواقع التي يرتبط بها هذا الموقع واتجاهاتها الفكرية ؟

اهتمت الدراسات المعنية بظاهرة الويب بتوفير معايير تقييم محتوى ومواقع الويب وقياس مستويات التحيز والمصداقية في المعلومات المقدمة على منصاتها. غير أن ذلك لا يعد معالجة حقيقية لظاهرة التحيز في مواد الويب، وإنما يوفر فقط رصدا أو قياسا لنوعية ودرجة هذا التحيز. وعلى ذلك فقد كان أحرى بهذه الدراسات أن تعمل على توفير الضوابط ومعايير البنية المعلوماتية والنشر الرقمي لمواد الويب بما يكفل اصدار مواد معلومات خالية من التحيز، بما يمكن أن يعلي مستوى تحقق ظاهرة مصداقية الويب **Web credibility** التي تذكر فقط كمقابل لجوانب التحيز وليس كمبدأ للنشر الرقمي لمواد معلومات الويب. ونظرا لاستمرارية سياسة القياس البعدي للتحيز في محتوى الويب بعد نشره، ظهرت أساليب مختلفة لرصد وقياس التحيز في محتوى الويب متمثلة في ؛ أولا أسلوب حكمة الحشود **Wisdom of crowds** ومؤداه أن يشترك مجموعة كبيرة من مستخدمي الويب في الحكم على محتوى مواقع الويب وتحديد جوانب ودرجات التحيز لهذا المحتوى، ثانيا نظرية تفسير الأهمية **Prominence-Interpretation theory** بحيث يمكن الحكم على طبيعة محتوى مواد المعلومات الويب وموضوعية المحتوى المتضمن[12]. وعلى ذلك فإن الاهتمام بظاهرة التحيز يختلف في متطلباته ومرحلته وإجراءاته عن الاهتمام بظاهرة المصداقية ؛ حيث يأتي التحيز في درجة متأخرة من حيث الفاعلية في التأثير على ضبط وتقنين المحتوى والمضمون الرقمي لمواد الويب، بينما تستطيع المصداقية إذا ما توافرت لها معايير وفعلتها أدوات النشر الرقمي على الويب أن تضمن سلامة المحتوى الرقمي لمادة المعلومات وإعلاء المصداقية على حساب اختفاء التحيز. ويؤثر كل من التحيز والمصداقية لمواد الويب على فاعلية وتفاعل نظم استرجاع الويب مع مواد المعلومات ؛ حيث يتمكن النظام من توفير درجة اكبر من الثقة ومعامل الشفافية في النتائج المقدمة إلى مستخدمي الويب. ولأن التحيز في

مواد الويب يختلف عن التحيز في أداء نظم استرجاع المعلومات، فإنه يجب مراعاة أن تكون نظم الاسترجاع نظما معلوماتية تبني وتصيغ استراتيجيات عملها على الموضوعية وعدم التحيز من جانبها[13].
يتباين التحيز في محتوى الويب على مستويين ؛ أولهما تدرج المسئولية عن المحتوى من المؤسسية إلى الفردية، حيث تتسم مواد المعلومات ذات النشر المؤسسي أو الرسمي بدرجة أكبر من الدقة في صياغة ونشر المحتوى والرسالة الفكرية بعيدا عن استخدام المصطلحات حادة التوجه التي قد تتبعها تاثيرات سلبية على هذه المؤسسات. بينما يغلب على المحتوى الشخصي الفردي لمستخدمي الويب من خلال منصات التعبير عن آرائهم مثل المدونات أو حسابات التواصل الشخصي أو حتى من خلال مواقع الاستضافة مدفوعة الأجر. ويمكن لنظم الاسترجاع التمييز بين هذن النوعين من التحيز من خلال رصد المصطلحات والمفردات النوعية الحادة التي غالبا ما تجميلها محركات البحث في قوائم الاستبعاد أو من خلال قدرات البحث الآمنة في مرحلة الاسترجاع. ثانيهما مستوى التحيز مواد المعلومات النوعية التي تختلف بين مواد تستهدف الرأي العام ومخاطبة مستخدمي الويب بشكل مباشر متمثلة في مواد إخبارية أو إعلامية تحمل بين طياتها رسالة موجهة لاستشارة تحفز معنوي أو تكوين رأي موجه لمواقف محددة، بينما تختلف طبيعة التحيز في المواد التجارية التي لا تحمل ضدية أو إساءة لطرف آخر مثل المواد الإعلامية، فهي مواد متحيزة لذاتها تسرد في ميزاتها دون التركيز على سلبيات الأطراف الأخرى كحالة المواد الإعلامية السياسية. وحري بالذكر أن الويب الإعلامية والسياسية تعود إلى خلفيات رسمية أو مؤسسية تضفي عليها قدرا من الدعم والتمكين، ومثل هذه المواد تحتاج من نظم استرجاع المعلومات دقة الإختيار وحذر في التحليل والمعالجة لهذه المواد فمن خلالها تتحول نظم الاسترجاع إلى أدوات نشر لتوجهات او أيديولوجيات مختلفة تصل مستخدمي الويب من نوافذ البحث بواجهات نظم الاسترجاع. أما مواد الويب العلمية فهي فئة معلوماتية يشترط فيها الحيادية والموضوعية، ومن ثم فإن قواعد بيانات الناشرين والدوريات المحكمة ومستودعات الأطروحات العلمية لا تجد نظم الاسترجاع صعوبة حيال الاختيار والمعالجة والبحث سوى توفير الدقة اللازمة في انتقاء الواصفات والمفردات الكشفية.

3 / 1 التوجهات النوعية وأشكال التحيز على الويب:

يرتبط التحيز اصطلاحا ومفهوم بوجود علاقة المقابلة أو التضاد بين توجهين ومن ثم يتشكل التخيز في دعم الرأي في مواجهة الراي المقابل ومحاولة تعزيزه بكالفة السبل. ويحمل التحيز المتضمن في مواد معلومات الويب مستويات عدة تختلف بين الندية والصدام في أعلى أشكاله وبين الدفاع عن وجهة النظر في أدنى مستوياته، مرورا بمستوى المنافسة أو المقابلة التي يحتاج كل طرف أو رأي تغليب توجهه على الراي أو الطرف الآخر، وهو ما يكون عادة داخل نطاق التنافس التجاري والإعلامي على الويب. وقد حملت مواد الويب خصائص هذه المستويات مجتمعة وفقا لطبيعة الدور ومجال التخصص المتضمن في محتوى مواد الويب ؛ حيث يحمل المحتوى الإعلامي ميزة تنافسية في قدرات التأثير على الرأي العام، بينما يحمل المحتوى السياسي في غالبيته الندية والتضاد مع التيارات

السياسية الأخرى، ومن ثم فأن مستوى ودرجة التحيز النوعي في مواد الويب يتأتى من طبيعة المجال والرسالة التي تحملها مادة المعلومات. وقد امتنعت نظم استرجاع الويب عن الدخول في هذه الدائرة بعدم تخصصها التجاري أو الإعلامي فضلا عن انتفاء التقسيم المبنى على توجهات أيديولوجية أو سياسية. وقد انحصر مبدأ التخصص والتقسيم لنظم الاسترجاع في التخصص النوعي أو الشكلي أو المعرفي لمواد معلومات الويب. غير أن نظم الاسترجاع عليها الحفاظ على هذه الحيادية في البناء والعمل في الوقت الذي تمتلئ قواعد بياناتها بمجموعات مواد المعلومات الموجهة والمتحيزة، وعلى ذلك ثمة أسلوبين بين مجاراة التحيز الكامن في محتوى مادة المعلومات وتغليبها في عمليات البحث والاسترجاع، أو الحيادية المطلقة للنظام مع هذه المواد.

1 / 3 / 1 التوجهات الايديولوجية:

تتدرج ألوان التحيز التي تصبغ مواد معلومات الويب بحسب طبيعة المحتوى الذي تتضمنه هذه المواد، حيث يظهر التحيز بأخف درجاته في المحتوى العقائدي أو الأيديولوجي برغم كونها مواد تحمل معلومات موجهة بطبيعتها، وتفسير ذلك أن المحتوى العقائدي أو الأيديولوجي سينصب على شرح أفكاره وتفسير جوانبه العقائدية والفكرية المخصصة دون الخوض في جوانب تخص نظرائها من العقائد أو الأيديولوجيات الأخرى، فنشر الذات الفكري هنا أهم من دحض الأفكار العقائدية أو الأيديولوجية المقابلة، ومن ثم فإنه لا يمكن نعت هذه المواد بالمتحيزة بل يكفي أن يطلق عليها مواد متخصصة في مجال عقائدي أو إطار أيديولوجي مخصص. وعلى ذلك فإن قاعدة قابتة يمكن البناء عليها تتمثل في أن التحيز لا يتمثل في المحتوى المخصص المتضمن لمجال ما، بل إنه محاولة بث أفكار ذات توجه ما متضمنة في محتوى يتناول موضوعا آخر. أما أكثر درجات ألوان التحيز فتتمثل في مواد المعلومات التي تحمل أفكارا وتوجهات مباشرة تتضمن المقابلة أو المناقشة دحض أفكار أخرى تتضاد معها. ويترتب على ذلك أن نشر النصوص المقدسة أو المتون الشارحة لها أو الكتب العقائدية والمحتويات الممثلة لأطر ودوائر عقائدية لا تعد جميعها موادا متحيزة، ومن ثم فإن تعامل نظم الاسترجاع مع هذه النوعية يكون وفقا لطلب مخصص واحتياجات معلوماتية مباشرة من جانب مستخدمي النظام، فلا طاقة لأي من نظم الاسترجاع أن يستدعي نصوصا عقائدية أو أيديولوجية مخصصة ضمن مجموعات النتائج المتصلة بمجال موضوعي ما. وهنا تضعف فرضية تحيز نظم الاسترجاع تجاه مواد الويب ذات المحتوى العقائدي سواء أكان ذلك التحيز متعمدا بتقديم أولوية الاستدعاء لهذه المواد، أم كان غير متعمد بإدراج هذه المواد خطئا في سياق نتائج مسترجعة، بواسطة تشابه الكلمات أو الواصفات الكشفية لهذه المواد مع مفردات الاحتياجات الموضوعية لمستخدمي النظام.

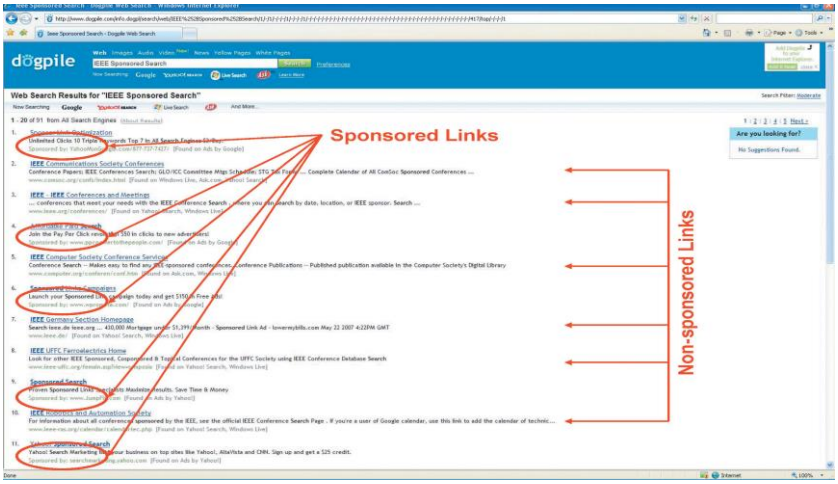
1 / 3 / 2 التوجهات الاقتصادية:

تمكنت الويب كبيئة للمطالعة والاستخدامي الإعلامي والمعرفي من التفوق على نظرائها من الصحف والتلفاز والإذاعة ؛ يستقي العالم الآن الأخبار من خلال البيئة الرقمية خلاف الإعتماد على أدوات

ووسائل الإعلام الأخرى. وليس هذا الأمر بغريب كون البيئة الرقمية الآن تحمل النسخة الرقمية من وسائل الإعلام والتواصل بين الناس، بما يجعل بيئة الويب بيئة رائجة وخصبة لنشر الدعاية والإعلانات التجارية التي تستخدف الوصول إلة أكبر عدد ممكن من الناس، وتحرض شركات الإعلانات على النشر في مواطن التجمع الأكثر مشاهدة لإعلاناتها. وقد غزت الإعلانات والترويج للسلع التجارية بيئة الويب ونظم استرجاع المعلومات على أنماط ثلاثة ؛ أولها إعتماد شركات الدعاية التجارية على استخدام مواقع الويب النوعية (الإخبارية والترفيهية ومواقع التواصل الإجتماعي) لنشر إعلاناتها حيث يرتادها العديد من مستخدمي الويب لأغراضهم المعلوماتية النوعية، وليس الأمر هنا إلا للمواقع الأكثر زيارة بما يجعلها أيضا الأوفر حظا في ترتيب النتائج المسترجعة داخل نظم استرجاع الويب. أما النمط الثاني للإعلانات التجارية فيتمثل في وضع الإعلانات التجارية على محركات البحث ذاتها والترويج للسلع والمنتجات من خلالها واتخاذها كمنصات للإعلانات التجارية من خلال العرض على صفحاتها المختلفة، وذلك دون التعرض لعمليات البحث أو التأثير على عرض النتائج المسترجعة لمستخدميها، وإنما حالها هنا كحال باقي المواقع التي تحصل على مقابل وضع الإعلانات على صفحاتها، وحري بالذكر أن رأس المال التجاري لشركات نظم الاسترجاع تعتمد على روافد الدعاية والاعلانات التجارية كمصدر للربح يصل إلى نسبة 95% من دخل نظم الاسترجاع، وقد أوضحت تقارير اقتصادية عدة أن حجم الإعلانات التجارية على محركات بحث الويب في العام 2007 وصل إلى 7 سبعة مليارات دولار، وهو ما يفسر تعاظم المنافسة المعلوماتية في خصائص البحث والاسترجاع لتنعكس على جلب أكثر عدد من مستخدمي الويب وبالتبعية تهافت الشركات الإعلانية والترويج التجاري على محركات البحث[14]. ولعل هذين النمطين السابقين هما الأقل في التعبير عن التحيز التجاري أو التوجهات الإقتصادية لمحركات البحث حيث لا تؤثر هذه الإعلانات الواردة في مواقع الويب أو المدرجة في صفحات محرك البحث على عملية البحث والاسترجاع أو التداخل مع النتائج المسترجعة والتأثير على ترتيبها في النتائج المستدعاة. النمط الثالث وهو التحيز الجلي لمحركات البحث المتمثل في بحث واسترجاع الإعلانات التجارية ذاتها ؛ حيث يتوجه مستخدمو الويب لمحركات البحث في الاستعلام أو البحث عن سلع ومنتجات تجارية يودون اقتنائها، ومن ثم يتعاظم مستوى التحيز التجاري ودرجته هنا في محركات البحث بأن تضع المحركات ترتيبا مسبقا أو منح شركات تجارية وسلعا بعينها ترتيبا أعلى ضمن قائمة النتائج المسترجعة، وهو ما يدعم توجيه مستخدمي النظام لمنتج محدد او سلعة بعينها، وبذلك يكون نظام الاسترجاع ذاته قد لعب دورا في الترويج التجاري باستخدام عملية البحث والاسترجاع ذاتها.وحري بالذكر هنا أن قضايا التجيز بالدعاية التجارية لنظم الاسترجاع تمثل تحديا أدبيا داخل نظم الاسترجاع التي تحصل على مقابل الدعاية والاعلان عن الهواتف النقالة الخلوية، بينما لديها في قواعد بياناتها عشرات الدراسات ومواد المعلومات التي تبنت التنبيه والشرح لمخاطر هذه الأجهزة والسلع على حياة مستخدميها[15].

لقد أتقنت نظم استرجاع الويب لعبة التعامل مع مواقع الويب في قائمة النتائج المسترجعة ؛ حيث تتاح مراتب الاسترجاع المتقدمة لمن يدفع مقابل الحصول والوصول لهذه المراتب. فلم تخفي بعض

نظم الاسترجاع ذلك بل وضعته تحت عنوان أولي في واجهة عرض النتائج يسمى **sponsored matches**، ويعرف هذا الإجراء في عملية الاسترجاع والترتيب داخل النظام بتحديد المستوى المدفوع مقدما **paid placement**. ومن ثم فالتحيز التجاري هنا ليس متعلقا بصفة موقع المعلومات، وإنما يتعلق بكيفية التعامل مع مواقع الويب والرسوم المتحصل عليها مقابل التدخل في ترتيب وتحديد مستوى العرض أمام المستخدمين. إن مثل هذا التحيز المدفوع مقدما لم يكن يسترعي انتباه مستخدمي محركات البحث ؛ حيث أجريت دراسة عام 2002 أثبتت أن واحد من كل أربعة مستخدمين يعون حقيقة النتائج مدفوعة المقابل، وعلى ذلك فإن الغالبية من مستخدمي نظم استرجاع الويب يتعاملون مع قائمة النتائج المسترجعة بثقة في موضوعيتها وعرضها وفقا لمدى اتصالها بالموضوع. وقد أشار 80% من مجتمع هذه الدراسة إلى وجوبية إعلام وتوضيح نظم الاسترجاع لمثل هذه الأمور التجارية لمستخدميه[16]، بل وتقنين التعامل مع المواقع ذات التعامل التجاري بتخصيص نافذة لهذه النتائج، وإلا فإن مستخدمي النظم يضعون مختلف النتائج في طبقة النتائج المضللة **deceptive advertising** كونها توضع عمدا بغرض الترويج وليس بغراض الإفادة والاتصال بالموضوع. ولم يقتصر التحيز التجاري لنظم الاسترداع على الحجز المسبق لمراكز متقدمة في النتائج المعروضة، وإنما امتد ذلك إلى عمل برناج الزاحف **spiders** في محرك البحث بحيث تصل إلى كشف صفحات ومستويات في مواقع الانترنت لم تكن لبرامج الزاحف قدرة على الوصول إليها أو التنقيب عن بيانات معمقة في قواعد بياناتها. ولطالما ظلت برامج الزاحف تعمل بشكل آلي ذاتي دون توجيه أو تعمد تحديد مسار ما لمواقع بعينها، وهو ما ثبت خطأه عند التعامل مع الموقع المدفوعة **sponsored** حيث تخلف درجة التكشيف ونمط التعامل معها دون غيرها[17].



شكل (2) النتائج مدفوعة المقابل للترويج في نتائج نظم الاسترجاع

لجأ مروجو السلع والمعلنون عنها إلى محركات بحث الويب نظرا للدور الإعلامي والإعلاني الذي اكتسبته محركات البحث لكونها واجهة الويب في التعامل مع المستخدمين، وذقد أكسبها ذلك قدرة كبيرة على الوصول إلى المستفيدين نظرا لوساطتها بين مواد الويب وتلبية احتياجاتهم المعلوماتية. ونظرا لهذا الدور فقد لعبت المحركات الدور الإعلاني والاتجاه نحو التربح بالترويج للسلع والمنتجات المادية أو الفكرية، وتطور هذا الدور إلى درجة التزام محركات البحث بالقيمة التجارية التسويقية على حساب القيمة المعلوماتية المنوطة بالحفاظ عليها في اختزان ومعالجة وبحث مواد المعلومات. ومن ثم فقد أدى ذلك الدور التجاري إلى خلق صفة التحيز الواضحة في استرجاع مواد المعلومات، وإصباغ محركات البحث بصبغة التمييز بين النتائج المسترجعة على حساب جوانب الترتيب وفقا للصلة الموضوعية ؛ حيث تكون الأولوية في عرض النتائج لمواد المعلومات المدعومة ترويجيا **sponsored materials**.[18]

وعملت نظم استرجاع الويب على استحداث تقنيات الإعلان والترويج مسايرة بذلك مختلف التطورات والمستحدثات التي طرأت على أجيال الويب الأول والثاني والثالث ؛ حيث بدأت نظم استرجاع الويب بالإعلانات الثابتة على صفحاتها حينما كانت الويب تعتمد في جيلها الأول على بنية المواقع غير التفاعلية التي تعتمد على تقديم المعلومات دون تدخل بالتعديل أو التحرير من جانب مستخدميها، وهو ما مثل الشكل الأولي لعملية الترويج التجاري في محركات البحث. وبالانتقال للجيل الثاني للويب حرصت نظم الاسترجاع على التفاعل مع مستخدمي الويب من خلال دمج المحركات والأدلة وتقديم خدمات للمستخدمين مثل إنشاء المدونات والصفحات الشخصية وقدرات التدوين والنشر الالكتروني الفردي، وهو ما تستخدمه نظم الاسترجاع كمنصات تجارية تطلق من خلالها حملات الترويج للسلع والمواد الإعلانية، فضلا عن ظهور التحيز التجاري حين إجراء عمليات البحث في قواعد بياناتها بحيث تجعل صفحاتها ومدوناتها وصفحات الخدمات الرقمية المتاحة عليها في مقدمة النتائج المسترجعة في عمليات البحث مضمنة خلالها المواد الإعلانية أو التجارية. على الجانب الآخر فإن قدرات النشر والتحرير الممنوحة لمستخدمي نظم الاسترجاع مكنت أصحاب التوجهات التجارية من أفراد وكيانات تسويقية من إنشاء الصفحات والحسابات الدعائية الإعلانية، بما يجعل صفحات نظم الاسترجاع التي يتم تكشيفها في قواعد بياناتها منصات إعلانية مجانية ذات قيمة استرجاعية داخل نظام الاسترجاع. وحري بالذكر أن نظم استرجاع الويب جرت في تيار التطوير لمستقر التحول من نظم استرجاع قائمة على معالجة مواد المعلومات إلى بوابات رقمية تقدم العديد من الخدمات المعلوماتية الرقمية وخدمات الحسابات وقنوات التواصل الاجتماعي فضلا عن دورها في استرجاع المعلومات، وهو ما زاد في قدرات الترويج والدعاية التجارية في مختلف الخدمات التي تقدمها بوابات نظم الاسترجاع[19].

شكل (3) الإعلانات الترويجية للسلع على بوابات نظم الاسترجاع

1 / 2 / 3 / 1 **التحيز لرعاة البحث** sponsored search bias

فعلت محركات البحث الشكل الإعلاني المتمثل في الرعاة الرسميين **sponsors** بتوفير قوائم مستقلة لمجموعات النتائج المسترجعة في البحث تخصص لهؤلاء الرعاة، وتقوم الفكرة هنا على أساس الدفع مقابل الزيارة **pay per click** وليس التعامل المبني على العرض وخلق انطباع لدى مستخدمي النظام أو لفت الانتباه فقط بالسلع المعن عنها ؛ حيث يتحدد مقابل الإعلان ورعاية البحث وفقا لمرات زيارة موقع المعلن أو الدخول على السلع المعلنة. وهذا الشكل أحد أسلوبين من أساليب رعاية البحث في نظم الاسترجاع، ولعله أقلهم تحيزا من جانب النظام للسلع أو المواد المعلنة في واجهة النتائج المسترجعة ؛ حيث يمثل استقلال القائمة المعروضة من النتائج وانفصالها عن قائمة النتائج الأساسية المسترجعة تنبها جليا من النظام لمستخدميه بشأن هذه النتائج وارتباطها بالجانب التجاري الإعلاني للنظام، فمن شاء فليستخدم أو يتجه بتركيزه إلى قائمة النتائج الأخرى. كما أنه يبتعد بنتائجه المسترجعة عن وصمة التحيز أو دعم الجانب التجاري على حساب الأهمية الموضوعية والصلة بموضوع البحث. أما الأسلوب الثاني وهو الأكثر ميلا للتحيز التجاري على حساب الحيادية الموضوعية هو الجمع بين نتائج الرعاة المدعومة تجاريا والنتائج المسترجعة وفقا للصلة بموضوع البحث، ولكونها تأخذ مكانا متقدما في العرض، فإن نتائج البحث المدعومة تستحوذ على انتباه واهتمام مستخدمي النظام أكثر من غيرها، وهو ما يرفع بزيادة عدد زياراتها وينعكس معه ازدياد الطلب والمقابل في عرضها من الرعاة. وعلى ذلك فإن ثمة اتجاهين لدى رعاة البحث أنفسهم للدعاية عن منتجاتهم يتمثلان في أولا : العرض المكاني وهو بمثابة مجموعات الإعلانات التي تستهدف حيزا جغرافيا تصل إلى دائرته، ثانيا : العرض السياقي وهو الإعلان المرتبط بطبيعة موضوع البحث أو مدى صلته بالنتائج المسترجعة المتوافقة والاحتياجات الموضوعية. وعلى ذلك فقد تعددت وتباينت أساليب العرض والتمثيل الإعلاني لمنتجات الرعاة على واجهات نظم الاسترجاع، وهو ما استتبع بدوره أن تتنوع آليات الترتيب والفرز **ranking mechanisms** داخل النظم على أشـــكال ثلاثة هي كالتالي [20]:-

أولا : نموذج الايرادات Revenue Model : وفيه يعمل النظام على تخصيص جزء من واجهة عرض النتائج المسترجعة قاصر على عرض نتائج رعاة البحث في النظام، بمل يجعل مستخدمي النظام على دراسة بالهوية التجارية الإعلانية لهذه النتائج. وتعرض النتائج في هذا الجزء بناء على الاسترجاع وفقا للمصطلحات المتصلة بموضوع البحث. ويحتسب نظام الاسترجاع إيرادات هذه النتائج إعتمادا على عدد مرات الزياة أو تنشيط هذه الروابط، ومن ثم فإن ايرادات النظام تتناسب وقدرته على الترويج لهذه النتائج.

ثانيا : آليات التوزيع Allocation Mechanisms :ويعتمد هذا الاسلوب على دمج النتائج التجارية داخل جملة النتائج المسترجعة المتصلة بموضوع البحث، ويكمن التحيز التجاري لنظم الاسترجاع في هذه الآليات ؛ حيث تمتزج النتائج المدعومة ماديا بالنتائج غير المروج لها، وهو ما يؤثر بالضرورة على ترتيب وعرض النتائج بينها وبين بعضها. ويضم هذا الأسلوب آليات عمل ثلاثة هي : -

1. تقديم العروض الأعلى سعرا **Highest Payers at the Top** :

وفيه تقوم آليات الترتيب والفرز اعتمادا على تاريخ الزيارات للمواقع المدعومة بحيث يتم الاسترجاع بناء على ارتباطها بالموضوع ثم يعاد التريب وفقا للأكثر زيارة حتى يتم تحصيل الموارد بناء على معدل الزيارة لهذه النتائج، ويتجلى التحيز التجاري هنا للنظام في تعمد تقديم نتائج على أخرى وفقا لتوقع الزيادة في معدل الزيارة وما يترتب عليه من زيادة إيراد الإعلان. غير أن هذه الآلية ترتبط بعوامل مثل الميزانية المخصصة لرعاية هذه النتائج، ومدى ارتباط الكلمات الدالة للنتائج المدعومة بكلمات البحث الموجهة للنظام، والتفاوت بين معدل الزيارة ومكان الترتيب لهذه النتائج في النتائج الإجمالية.

2. الجمع بين الصلة بالموضوع والعروض التجاري **Relevance and Bid Price Jointly** :

حيث تعمل آلية الفرز والترتيب اعتمادا على مدى صلة للموضوع وتوقع الاعتماد المالي المخصص لهذه النتائج، بحيث يعرض النظام النتائج المسترجعة إجمالا مع تقديم النتائج التجارية في اماكن متقدمة في قائمة الاسترجاع، ثم يتم احتساب الإيرادات لهذه النتائج وفقا لقرار مستخدمي النظام في زيارة أو إغفال هذه النتائج.

ثالثا : العرض المدفوع مقدما Posted-PriceMechanism:يعقد النظام اتفاقا تجاريا مع شركات الدعاية بحيث يحدد النظام مقابلا محددا لكل مكان او ترتيب ما في النتائج المسترجعة عند البحث عن موضوع أو مجال هذه المنتجات. ويقدر كل مكان في النتائج بمقابل محدد مسبقا وعلى النظام أن يوفر النتائج في أمكانها المتفق عليها في مرحلة الفرز والترتيب بحيث يكون مكانها بين النتائج وفقا للعائد المالي لهذه النتائج. وعلى ذلك فإن النظام يمارس التحيز التجاري في أوضح صوره لعرض النتائج المسترجعة وتفضيل المقابل المادي للعروض التجارية على خاصية مدى الارتباط بالموضوع لجملة النتائج المسترجعة.

1 / 3 / 3 التوجهات السياسية:

لقد اكتسبت نظم استرجاع الويب دورا إعلاميا يزداد مع تنامي حجم قاعد بيانات محركات البحث من مواد إعلامية وإخبارية تتضمن التعريف بالمعلومات حول قضية ما أو نقل أحداث بعينها، وقد جاء هذا الدور الإعلامي لكون محركات البحث النافذة البحثية عن المعلومات والقضاية الآنية التي

يستهدفها مستخدمو الويب بالبحث والمتابعة. ولم للدور الإعلامي لمحركات البحث أن يشكل أي أهمية في الحديث لولا ارتباط هذا الدور بتوجهات وقضاية قد يختلف تأويلها من رواية إعلامية إلى أخرى، ومن ثم فإن الرواية التي تتبناها محركات البحث أو تتحيز لها قد تكون الأكثر وصولا إلى مستخدمي الويب على انتشارهم العالمي والمحلي لمحركات البحث. بالاضافة إلى ذلك فإن الدور الإعلامي لمحركات البحث قد ارتبط بشدة بالمسار السياسي على المستويين المحلي والدولي ؛ حيث تداخل الدور الإعلامي لمحركات البحث مع الدور السياسي، على أثر نصرة رأي أو تيار أو حزب سياسي على حزب آخر، وهو ما ألبس محركات البحث ونظم الويب حلة الدور السياسي الذي حاد بمحركات البحث قليلا عن الدور المعلوماتي بالتحيز في عمليات المعالجة والبحث لنصر أيديولوجية أو فكرة ما على حساب الأخرى. وقد امتلكت نظم استرجاع الويب القوة المؤثر والفاعلة في تغطية الأحداث السياسية الدولية وخاصة أحداث النزاعات الأقليمية والقضايا السياسية الدولية وأحداث الانتخابات في الدول المختلفة على اختلاف درجات الأهمية بها. وامتلكت نظم الاسترجاع القدرة الفاعلة في التأثير على مستخدمي الويب وتوجيه تفكيرهم وانتباههم لصالح قضايا ما من خلال جملة النتائج المسترجعة أو الكلمات الدالة الموجهة التي تأخذ بيد امستخدم الويب لاتجاه دون آخر وفق التوجه والتحيز المدعوم من نظام الاسترجاع.

1 / 3 / 3 / 1 محركات البحث ولانتخابات الرئاسية:

وقد ظهر الدور السياسي لنظم استرجاع الويب من خلال الدور الذي لعبه محرك البحث **Google** في الانتخابات الرئاسية الأمريكية التي تزامنت مع إعداد هذه الدراسة بين مرشح الحزب الجمهوري دونالد ترامب ومرشحة الحزب الديموقراطي هيلاري كلينتون ؛ حيث بدا واضحا تحيز نظام البحث في جوجل لصالح هيلاري كينتون من خلال توجيه انتباه مستخدمي محرك البحث إلى مساوئ وسلبيات المرشح الجمهوري بأن ألحق **Google** الكلمات الضدية أو المسيئة لترامب ككلمات دالة مقترحة في عملية البحث إذا بدأ مستخدم **Google** في كتابة اسم ترامب، وعلى الجانب الآخر يقترح نظام البحث في جوجل كلمات دالة يمكن أن تضاف إلى اسم هيلاري كيلنتون وتأتي أغلبها بشكل إيجابي لتحسين صورة المرشحة الديموقراطية. ولعل ما انتهت إليه هذه الانتخابات بفوز الجمهوري ترامب وخسارة هيلاري كلينتون، رغم الدعم الكبير الذي حظيت به على محرك البحث **Google**، قد عازاه المحللون السياسيون إلى الدور العكسي الناتج عن دعم تام لمرشحة الحزب الديموقراطي والتحيز المفرط الذي جعل الديموقراطيين يركنون إلى الإطمئنان والتراخي عن الحشد والتعبئة لصالح هيلاري كلينتون مما كان أحد أسباب الخسارة للانتخابات الرئاسية. وقد حدث الأمر فعليا وفقا لدراسات أكاديمة أثبتت مدى تأثير محركات البحث والانترنت في مجموعات الناخبين الذين لم يكونو قد حددو تجاههم للتصويت لصالح أحد المرشحين جورج بوش الإبن أو آل جو في الانتخابات الرئاسية الأمريكية عام 2000، وقد أثبتت نتائج الدراسات أن الانترنت وجهت مجموعات أكبر لصالح آل جور بينما كانت وسائل الإعلام التقليدية تطرح جورج بوش[21]. وليس بعيدا ما حدث للرئيس الفرنسي ساركوزي الذي

طالبت فرنسا والاتحاد الأوروبي بحذف مواد معلومات تحمل الاساءة له والنيل منه من محرك البحث **Google**، وقد استجابت شركة جوجل لهذا الطلب ولم يعد إلا النادر من المعلومات المرتبطة بمساوئ الرئيس الفرنسي. وحري بالذكر أن الولايات المتحدة الأمريكية لا تمنع مثل هذه المعلومات أو التوجهات في محركات البحث انطلاقا من مبدأ تداول المعلومات وقانون حرية المعلومات. وقد قدم مركز ابحاث **pew research center** الواقع في واشنطن العاصمة بالولايات المتحدة الامريكية دراسة تكشف تنامي قدر الاعتماد على الانترنت كمصدر للمعلومات واستقاء الأخبار السياسية الخاصة باالانتخابات الأمريكية التي جرت في الفترة ما بين عامي 1992 و 2010، وقد كشفت هذه الدراسة تراجع دور الأدوات التقليدية أمام الانترنت ونظم استرجاعها. ويوضح الجدول (1) نسب استخدام الانترنت كمصدر للمعلومات في الانتخابات الرئاسية الأمريكية أمام الأدوات الإعلامية الأخرى[21]. وقد تتبع الباحث نسب الاستخدام في الانتخبات الرئاسية الأمريكية للعام 2016 التي اجريت وقت هذه الدراسة وقد عكست أن نسبة مستخدمي الانترنت لاستقاء المعلومات عن الانتخابات في الفئة العمرية من 18-29 عاما بلغت 35% و18% من المواقع الإخبارية مباشرة مقابل 12% من قنوات الأخبار التليفيزونية. وهو ما يعكس بدوره معدل النمو في استخدام الانترنت ونظم الاسترجاع في الانتخابات الأمريكية فترة بعد الأخرى[22].

جدول (1) نسب استخدام الانترنت كمصدر للمعلومات في الانتخابات الرئاسية الأمريكية

Presidential Elections					
	Television	**Newspaper**	**Radio**	**Magazine**	**Internet**
1992	82%	57%	12%	9%	---
1996	72%	60%	19%	11%	3%
2000	70%	39%	15%	4%	11%
2004	76%	46%	22%	1%	21%
2008	68%	33%	16%	3%	36%
Midterm Elections					
2002	66%	33%	13%	1%	7%
2006	69%	34%	17%	2%	15%
2010	67%	27%	14%	2%	24%

يستقي مستخدمو الويب معلوماتهم لتكوين آرائهم في مختلف المجالات من محركات البحث التي تشكل القاعدة المعرفية لطالبي المعلومات، وفي شأن مهم مثل الانتخابات الرئاسية الأمريكية بالنسبة لدولة متقدمة، فإن الناخبين في هذه الدولة يتوجهون إلى نظم استرجاع الويب لتشكيل وجهة نظرهم وتأكيدها في تجاه مرشح بعينه. وقد منحت الديموقراطية الأمريكية قدرا كبيرا من الأهمية لاستخدام محركات البحث في الجوانب السياسية ؛ حيث يكون فارق الفوز في الانتخابات غالبا بفارق ضيئل بين نسبة الفوز والخسارة لكلا المرشحين، وعلى ذلك فإن أعداد الناخبين غير المحددين أو الذين لم

يتخذو قراراهم النهائي تمثل هدف مهم جدا لكلا المرشحين بالاعتماد على استخدام توجيه نظم المعلومات وأجهزة الإعلام لكسب وحصد أصوات هذه الفئة. وتزداد أهمية استخدام محركات البحث في الشأن السياسي والانتخابات في الدول التي يقتصر استخدام محركات البحث فيها على محرك أو اثنين بدرجة أعلى من غيرها من نظم الاسترجاع، بما يجعل استمالة هذه المحركات وتحيزها لصالح مرشح ما هو فرصة للتأثير على أكبر عدد من الناخبين غير المحددين، ولتوجيه آراء الناخبين لصالح هذا المرشح من خلال صفحات النتائج والمعلومات المسترجعة الداعمة والمتحيزة له[23].

وقد أثبتت دراسة **Robert Epstein and Ronald**[24] أن هذا الدور السياسي لمحركات البحث أضحى عاملا مهما لترجيح العملية الانتخابية في الولايات المتحدة الأمريكية ؛ حيث أجريت هذه الدراسة على 4556 فردا، خرجت بنتائج أهمها أولا : أن تحيز الفرز والترتيب للنتائج في محركات البحث تستطيع التأثير على 20% أو أكثر من الناخبين غير المحددين **undecided voters** بتوجيه أصواتهم لصالح مرشح ما. ثانيا : أن نسبة التأثير على الناخبين جراء التحيز من قبل محركات البحث تزداد أكثر باختلاف الطبيعة الديموغرافية لمجوعات البحث. ثالثا أن دور التحيز في نتائج محركات البحث يزداد تأثيرا في الدول التي يستخدم أفرادها محرك أو أكثر بشكل محدد. وقد دفع هذا التأثير والتحيز في محركات البحث ان تنفق الشركات الأمريكية ما يقارب ال 20 مليار دولار مقابل أن تتصدر نتائجها قائمة النتائج المسترجعة لمحركات البحث، وذلك اعتمادا على دراسات المتابعة والتقييم لمستخدمي محركات البحث بأنهم يعمدون إلى تصفح نتائج البحث المسترجعة وفقا للترتيب والتنظيم الذي وردت به قائمة النتائج، ثم ما يلبثون أن يستخدمو النتائج الأولية التي أتت في مرتبة متقدمة، حتى وإن كانت النتائج التي أتت متأخرا هي نتائج الأكثر اتصالا والأقرب لموضوع البحث. وعلى ذلك فحري بلجان الدعاية والحملات الانتخابية في حجم الولايات المتحدة أن تنفق الملايين من الدولارات على استثمار التحيز وأولوية الترتيب في نتائج البحث، فضلا عن التحيز التام بدعم إيجابيات أو إظهار مساوئ المرشح الآخر. فقد أظهرت نتائج تقييم محركات البحث أنه من بين 300 مليون زيارة **click** للنتائج المسترجعة كان منها نسبة 92% للصفحة الأولى، كما أن النتيجة الأولى في النتائج المسترجعة حصلت على **32.5** % من اجمالي الزيارات والنتيجة الثانية **17.6**% من معدل الزيارات.

1 / 3 / 3 / 2 التحيز بمصادر الزمن الحقيقي real time search

فرضت تقنيات البحث في الزمن الحقيقي الوصول إلى مصادر معلومات نوعية تتعلق في معظمها بالأحداث الجارية، وأهمها مواد المعلومات الاعلامية التي غالبا ما تتناقل الأحداث السياسية والوقائع التي تتعلق بالأمور السياسية الجارية، التي يكون على رأسها في وقتها أحداث الانتخابات الرئاسية. ويمكن القول أن محركات البحث قد فرضت ذاتها على اللعبة السياسية في الدول المتقدمة، كونها الدول التي تتجه إلى استبدال مصادر المعلومات التقلدية بالويب ومحركات البحث، ومن ثم يتعاظم تأثير محركات البحث في التغيير الانتخابي وتعديل المشهد السياسي في هذه الدول. وقد ذكر **Mustafaraj, Eni**[25] أن 55% من الأمريكيين البالغين قد استخدمو الانرتنت في اتخاذ القرارات

السياسية بشأن الانتخابات الأمريكية في العام 2008، وهو بذلك معدل أعلى من العام 2004 الذي شهد 37% من الأمريكيين يستعينون بمحركات البحث في القضايا السياسية، وذلك للدلالة على ثقة الناخب الأمريكي في المعلومات المدققة من نتائج البحث. وقد استخدمت محركات البحث لإضافة أخرى تجعل المعلومات السياسية أكثر رواجا وتوافدا على نتائج البحث من خلال خاصية البحث في الوقت الحقيقي **real time search** التي تمكن مستخدمي محركات البحث من الوصول إلى مواد المعلومات الإخبارية ومواد مواقع التواصل الإجتماعي وعلى رأسها **facebook, twitter**. وهو ما يزيد من آنية وحداثة المعلومات السياسية المسترجعة في نتائج البحث وارتباطها بالتزامن الفعلي ومجاراة أحداث المشهد السياسي. وهذا يؤثر بدوره على تزايد قدرة نظم الاسترجاع في ضبط وتقنين توجهات الاخبار وتغيير مسارات الأحداث من حيث تقديم معلومات إخبارية على أخرى، بما يدعم تغليب أخبار ومعلومات على أخرى. وقد دشن نظام الاسترجاع **Google** تقنية البحث في الزمن الحقيقي عام 2001 وكانت ما تزال تحت التجريب حتى أضحت أهم بوابات البحث الإخبارية تتلقى ما يعادل **5,9** مليون زائر كل شهر، وذلك كونها تملك قدرة تكشيف **4,500** موقع إخباري باللغة الانجليزية وحدها، فضلا عن قدرة التحديث المستمر لهذه المقالات بشكل دوري كل 15 دقيقة[26].

لم تقتصر نظم الاسترجاع على لعب دور بوابات الأخبار فقط لمستخدميها في بحث المعلومات السياسية، إنما استهدفت النظم المتحيز التاثير على قطاع من مستخدميها ينتمي الى فئة عمرية مخصصة تتمثل في الشباب بين 18-25 عاما ممن لهم حق التصويت في الانتخابات والأحداث السياسية ؛ بحيث تمثل لهم النظم ليس فقط بوابات الأخبار وإنما بوابات المعلومات السياسية والروافد التي يستقون منها بنيتهم المعرفية حول الأشخاص والأحداث السياسية. وهذا الدور يتنامى بزيادة تنامي دور الانترنت واستقرارها كأسلوب حياة في هذه المرحلة العمرية. وحري بالذكر هنا أن خصائص الشخصية لهذه الفئة العمرية من مستخدمي نظم الاسترجاع تتسم بالتمرد على ادوات الاخبار والإعلام التقليدية كالصحف والتلفاز في متابعة أو الاخذ بما جاء فيها، ومن ثم فإن المصدر الرئيس هنا هو البيئة الرقمية المتحررة من قيود الدولة وتحجيم النشر أو تقنينه وتوجيهه. ولم يكتف مستخدمو نظم الاسترجاع بلعب دور المتلقي للمعلومات السياسية حول الأحداث الجارية، بل يملك الشباب قدرات النشر الإلكتروني وتحرير المدونات بما يجعله يقدم مواد معلومات تحمل وجهات نظر يعلوها التحيز الصارخ تبعا لطبيعة المفردات المستخدمة والتعبير عن الرأي، بما يجعل نظم استرجاع المعلومات قد تتبنى هذه المواد مرة أخرى في عملية البحث والاسترجاع، ومن ثم تتحيز مرتين بالتوجيه للقراءة وبالنشر والتداول لمنتجات معلوماتية بعينها[27]. وقد أظهرت دراسات المراكز البحثية للشئون السياسية في الولايات المتحدة بدراسة مدى تأثير المنشورات الداعمة للأفكار السياسية على مواقع التواصل الاجتماعي من جانب المستخدمين أنفسهم، التي أضحت في مرمى الاضافة والتكشيف داخل نظم استرجاع الويب. ومن ثم فإن هذه المواد الشخصية تمثل زخما وكما هائلا وصل في انتخابات 2008 إلى 113 مليون منشور يدعم التوجيه بانتخاب مرشح محدد وكانت في ضمن ما يمكن استرجاعه في محركات البحث عن الانتخابات الرئاسية الأمريكية. ونظرا لكون هذه المنشورات تأخذ

طابع التحيز فهي بطبيعتها ستكونسهلة التكشيف والتحليل من جانب مكشفات نظم الاسترجاع، بما يدعم مع هذا التحيز من الناشر التحيز القوي من نظم الاسترجاع في امكانية حجب او تقديم بعضها على الآخر[28].

لقد فرضت المنافسة السياسية الحزبية على الاحزاب نقل المعارك والمواجهات السياسية إلى بيئة الويب حيث الحدود الأوسع والأدوات الأكثر تأثيرا في مجموعات الشباب والمتابعين للأحداث السياسية. وتبلغ هذه المنافسات ذروتها في أثناء الحملات الانتخابية وأيام الاقتراع ؛ حيث قد تعرف كل أدوات الحشد والتعبئة صمتا انتخابيا إلا نظم الاسترجاع ومواقع التواصل الاجتماعي، فهي لا تعرف صمتا أو تستطيعه، ومن ثم تستمر حرب سياسية على الانترنت تحاول الاحزاب اقتناص المتابعين والفوز بها عن طريق الحشد والتعبئة للناخبين. وقد تبنت الاحزاب والحملات الانتخابية للمرشحين في الانتخابات الرئاسية والبرلمانية تدشين لجانا الكترونية تتولى مهام الحشد والترويج السياسي الالكتروني، وهو ما من شأنه التأثير على توجهات الناخبين والتأثير بأعداد المصوتين حتى اللحظات الأخيرة. وقد أثبتت الجولات الانتخابية في العديد من الدول وأهمها الولايات المتحدة الامريكية أن اللجان الالكترونية تدعم وتضغط لصالح مرشح أو توجه بعينه من خلال استخدام قدرات مواقع التواصل الاجتماعي ومحركات البحث ؛ حيث يمكن لمحركات البحث احتواء مختلف مواد المعلومات السياسية بغض النظر عن قيمتها واستخدام هذه المواد المتحيزة بطبيعتها لدعم توجه بذاته يتبناه نظام الاسترجاع. وتخرج محركات البحث بمنشورات التواصل الاجتماعي من محدودية قوائم الأصدقاء إلى آفاق البحث والنشر لمستخدمي نظم الاسترجاع وتمثيل موجة أعلى وأكثر امتدادا من نظام التواصل الاجتماعي[29].

وعلى ذلك فإن القيود التقنية لمواقع التواصل الاجتماعي لم تمنع نظم الاسترجاع من الوصول إلى نقل ومجاراة ووقائع الانتخابات بل وإدارتها والتحكم فيها عن بعد للتأثير على الناخبين، وهو ما يعطي قدرة أكبر لعمليات الضغط والتأثير من خلال التحيز المتعمد من جانب نظم الاسترجاع في انتقاء مواد المعلومات السياسية والإعلامية[30]. وقد تستخدم الدول متمثلة في النظم السياسية لها الجانب العكسي من عمليات التحيز في نظم الاسترجاع ؛ حيث أثرت زيارة الرئيس الصين **Jiang Zemin** لافتتاح مشروعات تابعة لوزارة صناعة المعلومات عام **2002** ، فقد أقدم أحد المسئولين على استخدام محرك البحث **Google** والبحث باسم رئيس الصين وفور كتابته في فراغ البحث اقترح نظام البحث مفردات حملت الاساءة للرئيس الصيني **"Evil Jiang Zemin"** مما جعل شركة جوجل تعمل على تغيير سياسية البحث والأداء في الصين تحيزا لصالح النظام السياسي في الصين[31]. ولقد استثمر الرئيس الأمريكي السابق أوباما المحتوى الرقمي وبثه من خلال نظم استرجاع الويب في صناعة ظاهرة اوباما **Obama Internet phenomenon** التي استشعرها الناخب الأمريكي في الانتخابات الرئاسية للعام 2008 ضد المرشح الجمهوري ماكين. وقد ساهم في إحداث هذا النصر قدرات نظم الاسترجاع على معالجة الوسائط المتعددة من مواد المعلومات السياسية المتحيزة لصالح أوباما واسترجاعها وبثها بشكل أوسع، بما حمل التأثير الموجه في الاتجاهين الايجابي لصالح المرشح الديموقراطي والسلبي

بشان المرشح الجمهوري، الذي لم يجد داعما له في قواعد بيانات محركات البحث على الويب بما يعني خسارة النتائج المرضية في محركات البحث قبل تحقق خسارة الأصوات في صناديق الاقتراع[32].

المبحث الثاني : بنية وتفعيل التحيز بنظم استرجاع المعلومات:

يعنى هذا المبحث بدراسة ملامح وجوانب التحيز داخل نظم استرجاع الويب وتركيبة الأداء المتحيز وعوامل تفعيله داخل جوانب الاسترجاع المختلفة التي تبدأ بالاضافة والاختيار والمعالجة والبحث أمام المستخدم. وذلك حتى تتضح الجوانب الإجرائية والعملية التي تحدث داخل نظام الاسترجاع ؛ حيث يمكن البناء على هذه الخصائص في قياس واقع التحيز داخل نظم استرجاع الويب في المبحث اللاحق.

2 / 1 الأداء المتحيز لنظم الاسترجاع :

ينقسم الحديث عن أداء نظم استرجاع المعلومات إلى شقين ؛ أولهما : يتمثل في اداء الويب كبيئة لاسترجاع المعلومات الرقمية ؛ حيث فرضت طبيعة الويب على مستخدميها أن يلجأو دوما إلى محركات البحث لاسترجاع وبحث مواد المعلومات النوعية شكلا وموضوعا. فما من سبيل للوصول والاتاحة لمواد المعلومات الا من خلال واجهات عرض النتائج في نظم استرجاع الويب، فمنها تكون مواد المعلومات ومنها يكون انتقاء الموضوعات للاستخدام والمعرفة. ويمكن أن يطلق على هذا الجانب الأداء الخارجي المرتبط ببيئة عمل محرك البحث مع مواد المعلومات واشتراط اتاحتها من خلال وسائط محددة، كما أن هذه البيئة ترتكن على النشر اكثر ما تهتم بالتنظيم والمعالجة المتروكة كليةً لنظم استرجاع الويب. وقد أشارت مختلف الاحصاءات ومؤشرات قياس الأداء جميعها الى تعاظم الأداء الخارجي لمحركات البحث في بيئة الويب حيث تمثل محركات **Google, Yahoo and MSN** أكثر أدوات الويب استخداما في عمليات البحث عن المعلومات، أيضا فقد أجرى الأمريكان وحدهم **6.4** مليات عملية بحث في العام 2006 فقط أو بما يعادل 210 مليون عملية لليوم الواحد، وقد جاءت نسب الاستخدام المحركات **12.8% and 28.5% ,44.7%** وفقا للترتيب السابق لمحركات البحث الأكثر استخداما. وقد زاد معدل هذا الاستخدام ليصل إلى 18 مليار عملية بحث في الشهر الواحد للعام 2015 في الولايات المتحدة الأمريكية، تركزت بنسبة 90% منها على محرك البحث **Google** بمفرده[33].

ثانيهما : يتمثل في الأداء الداخلي لنظم استرجاع الويب حيث رحلة مواد المعلومات داخلها البادئة بالاختيار والانتقاء ثم تحديد آليات المعالجة وسبل الاسترجاع، وهو ما يعول عليه أكثر في رسم سياسة الأداء المتحيز لمحركات البحث، حيث الانتقاء وفقا لمحتوى محدد والمعالجة وفقا لمفردات بعينها وتمييز وصفات على أخرى بحيث تساهم في الاسترجاع من زاوية محددة، والتركيز على رسم صورة معرفية لموضوع ما من خلال جملة النتائج المسترجعة التي تشملها قائمة عرض النتائج. وعلى ذلك فإن المحصلة النهائية لأداء محركات البحث أن تخرج مواد المعلومات التي تخدم توجه محدد دون الآخر أو تمييز لجانب من جوانب الحدث بشكل خاص.

2 / 2 آليات وأجراءات التحيز في نظم الاسترجاع:

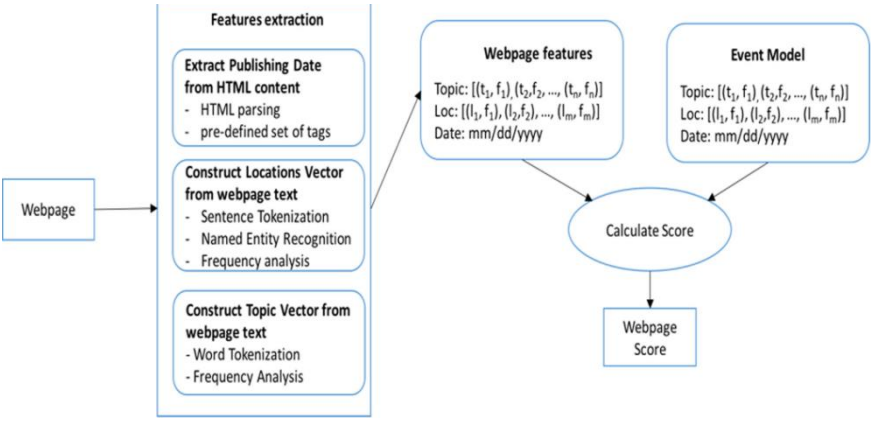
يرتبط الاداء الداخلي لنظم الاسترجاع بمجموعة من المكونات الداخلية يعمل كل منها في سياق وإطار مختلف عن غيرها من المكونات، ثم تتكامل جميعها في سياق عمل يمثل الاطار العام لعمل محرك البحث. وقد يكمن التحيز هنا إما في أحد هذه المكونات في ضبط أداء هذا المكون بتوجه وشكل محدد يؤدي إلى تغيير مسار العمل بشكل ما، أو أن يكمن التحيز في الأداء الجمعي لمكونات نظام الاسترجاع بحيث يترتب على ذلك التحكم المطلق في تدفق ومعالجة واسترجاع المعلومات وتحديد سياقات البحث ومسارات الاستدعاء. فكل من المكونات داخل النظام يملك بمفرده تغيير توجه العمل من المحايد إلى المتحيز بفعل التأثير المرحلي على دورة عمل النظام وتدفق مواد المعلومات داخل النظام ؛ حيث يمكن للنظام الفرعي لعملية التزويد والاضافة ضبط تحيز النظام على انتقاء واختيار مواد المعلومات من منبع الويب، وتشكل سياسات الاختيار والاضافة مجموعة التعليمات التي يتزود بها برنامج الزاحف، ومن ثم التحيز بإضافة مواد معلومات ذات واصفات بعينها والتوقف عن غيرها وفقا لسياسية التحيز بالاضافة. أما نظام التكشيف داخل كيان العمل فقد يتوقف عند مجموعات المصطلحات المحددة التي توافق تحيز النظام واختيار المفردات المعبرة عن رسم وجهة الاسترجاع وتوجيه مستخدمي النظام فيما بعد، حيث يتحكم نظام المعالجة بقدرات الربط بين مصطلحات البحث وهوية الموضوعات داخل مواد المعلومات. أما نظام البحث فهو الأداة الأخيرة التي تضبط وتقنن التحيز فيعمل وفقا لمنهج استرجاع واستخدام مصطلحات مقترحة يعني خروجها فقط امام المستخدم التأثير على قدرات اختيار مفردات البحث ومن ثم التشتت عن الهدف الأولي والتركيز على المصطلحات المتحيزة المقترحة[34].

2 / 2 / 1 الاضافة بالانتقاء:

تستخدم نظم بحث الويب برامج مخصصة للسعي خلف مواد الويب والبحث عنها اقتناءا وتزويدا لقواعد بياناتها، ومثل هذه البرامج التي تتمثل في الزاحف **crawler, spider** التابعة لمحركات البحث ترتبط بملفات **robot.txt** التي تحمل سياسة التزويد والاضافة لقاعدة بيانات محركات البحث. غير أن منهجية الأداء لهذه البرامج تعتمد على تحليل روابط صفحات الويب ومن ثم الانتقال منها إلى الصفحات المرتبطة بحثا عن مواد المعلومات النوعية المستهدفة، وتحدد تعليمات العمل في تتبع الروابط والكشف عن تحديثاتها أو حداثتها كلية وأخذ القرار بالاضافة أو عدمها لقاعدة البيانات. وقد تبنت نظم بحث الويب آليات متطورة تبني على اساسها برامج الزاحف تخالف آلية العمل السابقة، تنبني هذه الآلية على استخدام اسس الويب الدلالية والانطولوجيا بحيث يتم تتبع المحتوى وليس تتبع الروابط، ويتمثل هذا الأداء في تتبع مختلف مواد المعلومات النوعية التي تنتمي لموضوع واحد أو تلك التي تتعلق بحدث أو قضية واحدة. وعلى ذلك فإن ثمة تقنيات دلالية وبرمجية ترسم مسارات التحرك للزاحف يتوقف من خلالها عند مواد المعلومات المتصلة موضوعيا أو تلك التي تخدم حدث اجتماعي يمثل هدف مستخدمي الويب وموضع اهتمامهم. وبذلك تسعى نظم استرجاع الويب

لمسايرة موجات البحث واهتمامات مستخدمي الويب في تتبع معلومات تتعلق بأحداث سياسية كالانتخابات أو بأحداث اجتماعية أو بأحداث العنف والارهاب التي تشهدها الدول. واعتمادا على ذلك فقد انتقل الزاحف من الحصاد وفقا للروابط إلى الانتقاء وفقا للموضوع والحدث، وهو ما يؤكد امكانية تفعيل مبدأ التحيز وخدمة التوجه المحدد لنظام الاسترجاع ؛ بحيث يمكن للنظام من المنبع انتقاء والتقاط مجموعات مواد المعلومات التي ترسم وجهة نظر محددة أو لإقرار وضعا معرفيا اجتماعيا أو سياسيا بين مستخدمي نظام البحث. فقد ساهم تقدم تقنيات الأداء للزاحف زيادة قدراته على الغوص في محتويات مواد المعلومات وتحليلها وانتقاء موضوعاتها ومحتواها وفقا لما تخدمه هذه المواد، وتكاملا مع توجهات الزاحف في ضم مجموعات المعلومات التي تخدم توجهات النظام وتحيزه[35].

قدم **Mohamed M. G. Farag and Sunshin Lee**[36] في دراستهم نموذجا للزاحف المبني على تحليل النصوص لمواد معلومات الويب والتعرف على هويتها الموضوعية قبل إضافتها وانتقائها، وذلك من خلال تصنيف مواد المعلومات أولياً اعتمادا على نوع وموضوع مادة المعلومات **genre and topic**، ثم التوجه إلى تحليل النص المُضمَّن في مادة المعلومات والمحدد لمدى ملائمتها لتوجه الزاحف أم لا. ويقدم شكل (4) نموذج تحليل واقتباس ملامح مواد الويب النصية باستخدام محارف **tags** اللغة المعيارية للنصوص الفائقة **HTML** ومجموعات الميتاداتا **Metadata** المضمنة في مادة المعلومات لتحليل واشتقاق المفردات الدالة على الموضوع والمحددة والتغطية المكانية لمحتوى مادة المعلومات.



شكل (4) أداء الزاحف بانتقاء مواد الويب اعتمادا على مدى الصلة بالحدث والموضوع

يوضح شكل (4) أن برامج الزاحف تنطلق إلى إنتقاء مواد الويب محملة بمجموعات المفردات الدالة على هوية موضوعية وتوجهات محددة تمثل مرجعية الزاحف في انتقاء واضافة مواد معلومات الويب ؛ حيث يعمل الزاحف في مسارات ثلاث لاقتباس العناصر الموضوعية الدالة على هوية مادة المعلومات من حيث أولا التقاط الكلمات التي تدل على طبيعة الموضوع والتوجه المعرفي لمادة المعلومات وهو

المسار المعروف **topic vector** المهتم بالجوانب الموضوعية، ثم المسار الثاني المعروف **locations vector** المهتم التغطية المكانية والجغرافية لمادة المعلومات والمواقع المحددة للأحداث أو القضايا المثارة. ويتلو ذلك مسار المحدد الزمني وهو تاريخ النشر **publishing date** لمادة المعلومات وتتمثل أهميته في ارتباط الاحداث بفترات زمنية محددة أو الوقت الجاري لفاعليات الأحداث. وعلى ذلك يستخرج الزاحف محددات ثلاث هي الموضوع والمفردات الدالة عليه، والمكان والمفردات المحددة له، والتوقيات الزمنية والتواريخ المرتبطة بها[37]. وتمثل هذه المحددات إجمالا الصورة العامة لتحليل مضمون مادة المعلومات وتقدير الوزن النسبي لأهميتها باستخدام أوزان المفردات التي وردت بها، وضمها إلى قاعدة البيانات. ويحتاج الزاحف لاتمام هذا العمل إلى مقومين أساسيين هما القدرة الدلالية على الربط بين المفردات المختزنة الدالة على التوجه والبين المفردات المختلفة الواردة في مادة المعلومات، أيضا تقدير أهمية مادة المعلومات ومنحها وزنا نسبيا من درجات ثلاث هي **A, B and C** يمكن أن تتحرك به في في رحلتها من الإضافة إلى الاسترجاع أمام مستخدمي النظام. ويمثل ذلك تقدما نوعيا في قدرة الزاحف على جمع مواد المعلومات الممثلة لرأي أو توجه واحد وتتشارك جميعها في مضمون وأفكار متماثلة ومتوافقة، بما يعزز الاداء المتحيز لنظام الاسترجاع في دفع مواد معلومات بذاتها إلى مستخدميه موجها تركيزهم إلى ما يتحيز إليه النظام. وقدم الباحثان نموذجا تطبيقيا لعمل الزاحف في جمع مواد الويب حول أحداث محددة مثل أحداث إطلاق النار في ولاية كاليفورنيا **Bernardino, California, on December 2, 2015** ؛ بحيث يتم انتقاء مواد المعلومات التي تحتوي على مفردات مختلفة تمثل الموضوع بجانب مفردات تحديد موقع ومكان الحدث ثم المحددات الزمنية لمواد المعلومات التي نشرت في هذا الوقت، وكانت هذه المسارات متكاملة في المثال التالي :-[38]

– Topic: shooting, shooter,. . ., etc.

– Location: San Bernardino, California

– Date: 12/02/2015

واعتمادا على هذه الأداء فإن التحيز في أداء الإضافة والتزويد يتم وفقا للخطوات التالية :-

1. تحديد قائمة الأحداث أو الأشخاص أو الوقائع موضع الاهتمام للنظام.
2. ضبط قائمة استنادية تتوافق وبنية الانطولوجيا والويب الدلالية لمفردات كل حدث او موضوع.
3. تزويد برنامج الزاحف بمختلف المفردات التي تدعم الإضافة أو الرفض لمواد المعلومات.
4. تحليل صفحات الويب من حيث مفردات الموضوع والمكان والتاريخ.
5. تحديث قاعدة بيانات النظام بمفردات الموضوعات والأحداث والاضافة تبعا لتغير المفردات الدالة
6. تحديد درجة أهمية صفحات ومواد المعلومات وفقا لتحليل محتواها في المستويات **A, B, C**.
7. ضم صفحة الويب واتخاذها كجسر للمرور للصفحات المرتبطة معها المتشاركة في الحدث.

2 / 2 / 2 **التكشيف** indexing algorithm

غيرت أدوات الإضافة والتكشيف في نظم الاسترجاع سياستها في التعامل مع عمليات الاختيار وتحليل المحتوى لمصادر الويب من الاعتماد على المواقع المترابطة برمجيا إلى مجموعات الصفحات المترابطة موضوعيا، بما يعني التركيز محتوى مادة المعلومات أكثر من عشوائية التحرك عبر روابط المواقع. فتتّبع الآن آليات التكشيف في نظم الويب النظام العنقودي في التحرك بين صفحات الويب عبر تحليل اسم المجال **domain name** بحيث يمكن للزاحف التحرك وفقا للمحتوى وتشعباته أكثر من الاعتماد على الروابط الداخلية لمحتوى الموقع التي تتسم بعشوائية التوزيع الموضوعي. وعلى أثر ذلك فإنه يسهل على نظام الاسترجاع الآن تتبع موضوع بذاته أكثر من تتبع المواقع، ومن ثم فإن تكشيف مواد الويب النصية والوسائط المتعددة يعتمد على مجموعات الاشارات النصية من ميتاداتا **metadata** وغيرها من محارف اللغة المعيارية التي تحمل هوية وموضوع مادة الويب[39]. وتكمن أهمية التكشيف كأداة تحليل للمحتوى في نظام الاسترجاع كونها تحدد فقط الاستخدامات الموضوعية والتوجهات الفكرية التي ستخدمها مواد المعلومات في قاعدة بيانات النظام ؛ حيث لا يتم البحث والوصول والاتاحة لجملة مواد الويب داخل النظام إلا من خلال الكلمات الدالة الكشفية التي يمنحها المكشف **Indexer** لكل مادة بشكل مستقل، وعلى الوجه الآخر من عُملة التكشيف تستقر الكلمات المفتاحية التي ستحقق فاعلية المضاهاة **matching** بين الكلمات الدالة من جانب المكشف والكلمات المفتاحية من جانب مستخدمي النظام. ولذلك فإن حصر استخدام أو انسياب تدفق مواد المعلومات يحدد بجانب آلية التكشيف المتبعة من جانب النظام وليس من خلال محتوى مادة المعلومات. فالمكشف يستطيع تحويل وجهات الانتماء الموضوعي وخدمة عمليات بحث مخصصة من خلال التحول بين آليات التكشيف واستخدامها بالتناوب على محتوى مادة الويب، فثمة آليات قد ترفع مصطلحات على أخرى في محتوى مادة المعلومات وهو ما يجعل دلالة هذه المصطلحات هو الموضوع الغالب على محتوى الوثيقة، وهكذا في حال رفع درجة مصطلحات مختلفة أو مغايرة للمصطلحات السابقة، ومن ثم يتلون محتوى مادة المعلومات بتلون المصطلحات الكشفية الناتجة عن برنامج المكشف[40].

وقد استخدمت نظم الاسترجاع لعمليات التكشيف أكثر من نمط في الوقوف على طبيعة وهوية موضوع مادة المعلومات وربطها بمستوى محدد أو نوعية بذاتها من التوجهات الموضوعية، وهذا بدوره يعزز من قدرات التحيز وتفعيله داخل النظام، بل والتحكم فيها بتوجيهها في صالح عمليات بحث مخصصة أو حجبها عن الاستخدام في أوقات محددة، وذلك كله من خلال قدرات التحكم في طبيعة المصطلحات الكشفية المميزة والمعبرة عن مادة المعلومات. فالتحول بين آلية التكشيف بالاحصائية إلى التكشيف باللغة الطبيعية إلى التكشيف الدلالي يجعل مادة المعلومات تنتقل من خدمة موضوع إلى آخر أو أن تختلف على الحد الأدنى من درجة ما في خدمة موضوع إلى درجة أعلى أو أقل لنفس الموضوع باختلاف المصطلحات الكشفية التي الناتجة عن عمليات التكشيف. وقد كانت آليات التكشيف تُبنى بالأساس على نظام التكشيف المقلوب من صفحات الويب، وهنا يتم النظر إلى المعاني والدلالات الظاهرية للمصطلحات الدلالية، وبالتالي يتم التوجه إلى الاعتماد على اقتباس

الملامح الموضوعية من مجموعات ميتاداتا العناصر والكائنات مثل الوسائط المتعددة التي توجد داخل الموقع والصفحة ذاتها وليس نص الصفحة والصياغة المتحيزة، التي يمكن أن تصاغ من الأساس على تحيز ما من منشئ صفحة الويب وومحتواها المضمن داخلها. وحري بالذكر أن آلية التكشيف المتبعة من جانب النظام تمثل في ذاتها مقياس لدرجة التحيز التي يتسم بها النظام في تقديم أو تأخير جوانب موضوعية محددة بتقديم وتأخير المصطلحات الموضوعية الدالة عليها[41].

يختلف مستوى التحيز في عمليات التكشيف داخل النظام باختلاف اسلوب التكشيف المتبع مع مواد الويب ؛ حيث يعتمد النظام أحد اسلوبي التكشيف الاحصائي أو التتكشيف المفاهيمي **statistical indexing, concept indexing**، ويعتمد التكشيف الاحصائي على استخدام مبادئ الاحصاء والتكرار والتعامل مع وحدة الكلمة ضمن مجموعات الوحدات الواردة في مادة المعلومات كوحدة مستقلة دون الاهتمام بالسياق أو المضمون التي وردت به هذه الكلمة. وعلى ذلك فالاحتمالات الأعلى للكلمات الدالة على الموضوع تكمن في الكلمات المرجحة احصائيا عن غيرها من كلمات السياق لما لها من درجة تكرار مرتفعة، وقد فطن منشئو مواد الويب ومصمموها إلى هذا الأداء من جانب نظم البحث، فاستخدم محررو مواد الويب بعض الكلمات بشكل متكرر مع تضمينها في أجزاء محددة من مادة المعلومات يهتم بها المكشف عند اختيار الكلمات الدالة للمحتوى. أما التكشيف المفاهيمي فيعتمد على التعمق في محتوى الكلمة والنظر إلى سياقها كرؤية شاملة دون التركيز على وحدوية الكلمة، ومن ثم فإن الكلمات الدالة أو الكشفية المعبرة عن مادة المعلومات ليست تلك الأكثر تكرارا، بل تلك التي تعبر عن سياق الوثيقة ومحتواها من خلال دلالات العبارات ومادة المعلومات ككل[42].

ويساهم أسلوب التكشيف الاحصائي دعم قدرات التحيز لنظام الاسترجاع حيث يستطيع النظام التعامل احصائيا مع مادة المعلومات لاختيار كلمات محددة بتطويع خوارزميات الاختيار والتكشيف التي تضبط أدائها وفقا لتوجه محدد، وذلك لكون الكلمات وحدات مستقلة متراكبة لتكوين السياق العام، أما التكشيف المفاهيمي فيضع نظام الاسترجاع في مأزق التجانس والسياق أكثر من الوقوف عند مستوى الكلمات المتناثرة، بما يجعل النظام مقيد بمحتوى وموضوع مادة المعلومات أكثر من ارتباطه بالكلمات المفردة. وتتشكل لدى النظام هنا شبكة من المفاهيم المترابطة على مستوى مادة المعلومات ترتبط بعلاقات موضوعية ومنطقية ترسم الملامح الموضوعية لمادة المعلومات كاملة، ومن ثم فإن العناصر الدالة لا تتمثل في الألفاظ والكلمات المشتتة، بل في المفاهيم التي تعبر عن أفكار وموضوعات مادة المعلومات[43].

3 / 2 / 2 الإكمال التلقائي auto complete algorithm

تستخدم نظم استرجاع الويب مختلف آليات الوصول والبحث أمام مستخدميها بهدف تيسير اتاحة مواد معلوماتها، ومن بين هذه الآليات جاءت تقنية الإكمال التلقائي **auto complete** لدعم المستفيد في اختيار مصطلحات البحث التي تنطوي على مفاهيم متخصصة أو أطر موضوعية محددة. وتنطوي هذه التقنية على جعل المستفيد أمام خيارات متعددة لاختيار كلمات البحث، ففور

أن يشرع مستخدم النظام في كتابة مصطلحات البحث يعمل النظام على تزويده بمقترحات **Suggestions** لاكمال السياق الذي بدأه المستخدم. وتعود هذه التقنية إلى العام 2008 بعد شيوع استخدامها مع **Google**، وقد بدأت ارهاصاتها منذ 2004 حين بدأت نظم الويب في دعم هذه التقنية لدعم مستخدميها. ويمكن في هذه التقنية تدخل النظام في التأثير على اتجاهات البحث ؛ حيث تعمل مقترحات النظام على الزام المستفيد بمسارات موضوعية وتوجهات بحثية محددة قد يأخذ بها المستخدم أو يتركها تماما إذا لم توافق هذه المقترحات احتياجات المستخدم من مواد المعلومات. وتمثل مقترحات النظام لمجموعات المصطلحات والاختيارات المطروحة أمام المستخدم تحيزا من جانب النظام لمسارات بحثية محددة يختارها النظام، وهي بالتالي قد تساهم في توجيه المستخدم إلى مصطلحات لم تكن في نطاق اهتمامه أو التأثير على اتجاهات بحث المعلومات من خلال مجموعات المفردات المتدخلة في سياق البحث. وبرغم كون هذه التقنية تهدف لدعم المستخدم وتزويده بمقترحات لصالح البحث، غير أنها تمثل تدخلا متعمدا من جانب النظام أو على الدرجة الأقل نوعا من التشويش على المستخدم في تحديد هوية مفاهيم ومصطلحات البحث. ولم تكن هذه الآلية لتمثل شكلا من أشكال التحيز من النظام والتدخل في عملية البحث لو أنها تسير بشكل آلي وفقا لمعطيات برمجية أو تسييرا مبرمجا ينبع من طبيعة استخدام آليات البحث في النظام، غير أن هذه الآلية تخضع لسيطرة وتحكم النظام بشكل عام، فضلا عن قدرات التحرير والتعديل في هذه المقترحات بالاضافة والحذف وإعادة ترتيب عرضها أمام المستخدم، بما يعني قدرات أكبر للنظام للتدخل في اختيار المصطلحات البحثية وما يترتب عليها من تغيير في استراتيجيات البحث ينعكس بدوره على النتائج المسترجعة النهائية أمام المستخدم[44].

تنطلق آلية الاكمال التلقائي في النظام اعتمادا على جوانب البيئة واللغة وسمات المستفيد ؛ حيث تستقي هذه الآلية مفرداتها المقترحة على المستفيد من عمليات البحث التي تمت باستخدام النظام في إطار المنطقة الجغرافية التي ينتمي لها المستفيد، أو عمليات البحث السابقة التي أجراها هذا المستفيد أو من خلال لغة البحث المستخدمة في مواد المعلومات. ووفقا لما ينتهجه نظام البحث **Google** من تخصيص البحث على المناطق الجغرافية واللغات الانسانية والأديان، فإن المفردات المقترحة تتغير بطبيعتها من نطاق بحث إلى آخر وفقا لهذه المتغيرات[45]. غير أن ذلك لا يمنع تدخل النظام في تمرير أو حجب بعض الكلمات التي يمكن تعمد إدخالها أو التوقف عنها أثناء البحث، وقد استخدمت فرنسا من قبل حق التدخل القانوني لمنع **Google** من استخدام مفردة **arnaque** التي تعني بالفرنسية احتيال، وذلك بعد أن اقترن اقتراحها بهيئة التعليم عن بعد مما عاد على هذه الشركة بالضرر المباشر والترويج السلبي لهذه الهيئة. كما أن نظم الويب تستطيع التدخل في إعادة ترتيب أولويات البحث أمام المستفيد وليس فقط تغيير المفردات ذاتها ؛ حيث يتدخل النظام في ترتيب ووزن أهمية العبارات المعروضة وفقا لما يراه النظام أكثر صلة بالموضوع دون الالتزام بأكثر المصطلحات بحثا أو اكثرها استخداما في الموضوع، مما يدل بشكل قاطع على أن هذه الآلية تمثل وجهة نظر النظام في اجراء عمليات البحث، بل وتعمد التدخل في اختيار المصطلحات وسياق البحث بما ينعكس على

النتائج المسترجعة. وتتكامل هذه الآلية مع آلية التدقيق والتصحيح اللغوي بحيث تجرى عمليات البحث وفقا لمصطلحات صحيحة من البداية بما يحافظ على دقة ووقت البحث لمستخدم النظام. وحري بالذكر أن آلية الاكمال التلقائي لا تعنى فقط بالاقتراح، بل بالمنع أيضا حيث تتدخل نظم البحث في التوقف عن استخدام مفردات محددة تتعلق بالجوانب القانونية للهيئات أو الترويج السلبي للمنظمات أو المفردات المتعلقة بالأمن والكراهية والعنف او تلك المتصلة بالمعلومات السرية للأشخاص وسرية بياناتهم الشخصية[46].

شكل (5) شكل تباين المقترحات بين كلمتي البحث على نظم الاسترجاع47

4 / 2 / 2 البحث والاسترجاع search algorithm

تمثل آلية البحث الشكل النهائي الذي يصاغ عليه استفسار البحث وشكل الربط بين مجموعات الكلمات المفتاحية المستخدمة للتعبير عن الاحتياجات المعلوماتية للمستخدم. ويكون المستخدم هنا هو مصمم ومنفذ استراتيجية البحث بما يتوافر له من آليات ومعاملات بحث متاحة من جانب النظام، فاستخدام آليات الربط بالمنطق البولياني أو آليات التقارب بين الكلمات أو معاملات البتر المختلفة. وينعكس شكل استراتيجيات البحث النهائية على النتائج المسترجعة من النظام، بل ويحدد شكل استراتيجية البحث الوزن النسبي للكلمات المستخدمة وترتيب أولويات النتائج في عرضها. وتمثل مرحلة اجراء واختيار استراتيجية البحث المرحلة الوحيدة التي لا يتدخل النظام في أدائها ؛ حيث يملك المستخدم فقط اختيار الكلمات النهائية للبحث ومعاملات الربط بينها وفقا لما يراه مطابقا للتعبير عن احتياجاته، دون أن يتدخل النظام بالمنع أو التغيير في الكلمات والمعاملات المستخدمة. وذلك ليس معناه عجز النظام عن التحيز في هذه المرحلة، بل إن التحيز يأتي سابقا ولاحقا لهذه الخطوة ؛ سابقا في مرحلة التكشيف واختيار الكلمات الدالة المعبرة عن المحتوى أو منع كلمات بذاتها من التعبير عن محتوى مادة المعلومات، ومن ثم يترتب على ذلك التوقف عن التعامل مع مفردات بذاتها حتى وإن اشتملت عليها استراتيجية البحث المقدمة من جانب المستخدم، أما لاحقا فيكون عن طريق اجراءات

الترتيب والفرز التي يمكن أن يتدخل فيها النظام بتقديم مواد على أخرى أو بحجب مجموعات المواد التي تحتوي على مصطلحات محددة. فالبحث هو اسهام المستخدم في استكمال دورة الاسترجاع داخل النظام وليس للنظام طاقة بالتدخل الاجباري في اختيار أو الغاء شكل ما لاستراتيجيات البحث المقدمة له، أما باقي الاجراءات والادوات فهي تمثل وجهة نظر النظام في إدارة مواد معلومات الويب في قاعدة البيانات الداخلية[48].

يبدأ تحيز النظام في عملية البحث وتغليب وجهة نظره مع بداية رسم سياسة التكشيف وتنفيذ أولى خطواتها ؛ حيث تنعكس سياسة التكشيف ولغاته على استراتيجيات البحث والشكل النهائي لها، وهو ما يرسم مسبقا اتجاه البحث والآليات المستخدمة لتحقيقه. فثمة فارق نوعي في التعامل مع اللغات المقيدة والطبيعية في البحث التي تمثل بطبيعتها منهج التكشيف داخل النظام. فاللغة المقيدة هي لغة النظام وليست لغة مواد المعلومات بما يعني أنها لغة انتقائية للمفردات وتشكيل تام لتوجه النظام في التعبير عن إطار موضوعي أو فكري محدد، ويتجلى ذلك في إملاء النظام على مستخدمه التقيد بهذه المصطلحات وعدم الخروج عن مساراتها أو سياقها. أما اللغة الطبيعية فهي لغة مواد المعلومات القريبة من المؤلف والمستخدم البعيدة عن النظام وتوجهاته، وفي نظام إدارة المعلومات فإن النظام عليه أن يكون جسرا محايدا بين مصدر المادة المعلوماتية ومستخدمها وليس وسيطا متدخلا متحيزا بالحجب أو الاقتراح. ونظرا لطبيعة الويب وتدفق المعلومات بها باتت النظم عاجزة عن تفعيل التحيز بالقدر المطلوب باستخدام اللغة المقيدة وأضحت اللغة الطبيعية هي لغة التكشيف داخل نظم استرجاع الويب، فللغة الطبيعية تملك مرونة وانسياب المفردات التي تفرض بها التعبير عن محتوى المادة بختلف المفردات الممكنة، ومن ثم فليس من نظام للغة المضبوطة يمكنه السيطرة على مختلف هذه المفردات وشتات معانيها ودلالاتها. وينعكس الاختيار بين لغتي التكشيف على أساليب البحث المستخدمة في النظام ؛ فاللغة المقيدة تنتج أسلوب التصفح المقيد بالاختيار بين المصطلحات،. أما اللغة الطبيعية تنتج أسلوب البحث الحر المرتبط باختيار الكلمات والتعبير عن الاحتياجات المعلومات بكامل حرية مستخدم النظام. وبين هذين الأسلوبين بون شاسع في درجتي التحيز والتدخل من نظام الاسترجاع، فأعلاها في تقييد المستخدم بمصطلحات التصفح وأدناها السماح للمستخدم بالتعبير عن احتياجاته المعلوماتية وفقا لما يروق له من مفردات ومعاملات بحث متنوعة[49].

تتوزع آليات البحث في نظم الويب بين ثلاثة مستويات نوعية تمثل الشاكلة العامة لأداء نظم الاسترجاع ؛ حيث تعكس آليات البحث شكل العمليات السابقة من تكشيف وتحليل داخل قاعدة البيانات، وكيفما كانت أنماط التكشيف وآليات اختيار المصطلحات الدالة على هوية مادة المعلومات، كيفما كانت آليات البحث المتاحة امام مستخدم النظام. وتتحد آليات الاسترجاع والبحث في أنماط ثلاث تتدرج في التعامل مع المفردات من درجة الشكل واللفظ مرورا بدرجة التحليل اللغوي والتركيبي وصولا إلى درجة الدلالة والمفهوم لكل مصطلح على حده. وحري بالذكر أن انما التعامل الثلاث هذه التي ستنبني عليها آليات الاسترجاع هي التي تمثل بادائها مدى وقدرة تدخل النظام في توجيه النتائج

والتحكم في مواد المعلومات المستدعاة إلى المستخدم ؛ حيث يكمن أعلى مستويات التحيز في أقل درجات الأداء وهو التعامل مع المصطلحات والمفردات وفقا للشكل واللفظ، أما أدنى درجات التحيز فتكمن في أعلى مستويات العمل التي تتعمق في الدلالة والمفهوم المتضمن في شكل وحروف المصطلح.

وعلى ذلك فإن أنواع التحيز في أداء البحث والاسترجاع داخل نظم الويب تتمثل في الأنواع التالية :-

أولا : البحث المنطق البوليني : يستخدم البحث بالمنطق البوليني معاملات الربط بين مجموعة مفردات البحث وفقا للوزن النسبي لهذه المفردات وأهميتها في التعبير عن الحاجات المعلوماتية للمستخدم. وتمثل المعاملات الثلاث للمنطق البوليني الصياغة العامة لاستراتيجية البحث ؛ حيث تلزم أداة الربط **and** باستدعاء مفردات خاصة تمثل أعلى درجات الأهمية في طلب هذه المصطلحات، أما أداة البحث **or** فتعبر عن امكانية التباديل والتغيير بين المصطلحات بحيث يمكن استدعاء مواد المعلومات لموضوع ما وفقا لمختلف الجوانب أو الأوجه التي تناولت الموضوع، أما أداة النفي **not** فتضع القيمة الصفرية الاجبارية لمجموعة المصطلحات السلبية في البحث. إلا أن مختلف معاملات المنطق البوليني تتعامل مع الاطار الخارجي للمفهوم المتمثل في شكل ولفظ البحث بحيث يحتكم إلى مبدأ التطابق أو المضاهاة بين المصطلحات البحثية في استراتيجية البحث والكلمات الدالة الكشفية داخل مادة المعلومات، دون الامتداد إلى المعنى أو دلالة المصطلح. وعلى ذلك فيمكن للنظام أن يتدخل متحيزا بمنع أو تمرير مصطلحات أو ألفاظ محددة في استراتيجية البحث ومن قبلها في الكلمات الدالة المعبرة عن محتوى مادة المعلومات، ويسهل ذلك تفعيله من خلال قائمتي الحجب والتمرير **stop / go list**. فالبحث هنا يتم وفقا للفظ وشكل الكلمة هو ما يسهل التحكم فيه من جانب النظام، واطلاق العنان لقدرة التحيز والتدخل ضد أو مع اتجاهات موضوعية وفكرية محددة، وبناء عليه فالبحث هنا يمثل أعلى درجات التحيز للنظام في الاسترجاع والبحث[50].

ثانيا : البحث باللغة الطبيعية : يتقدم مستوى البحث باللغة الطبيعية عن البحث بالمنطق البوليني في قدرة التحليل والتركيب لمجموعة مفردات البحث ؛ حيث تتمكن آليات البحث باللغة الطبيعية من استخدام أدوات عمل مثل المعاجم اللغوية والمحللات الصرفية تعمل بدورها على تحليل مفردة البحث وتغيير طبيعة المفردة ودلالتها بتغيير شكلها وبنيتها اللغوية بما يساعد على توسيع إطار البحث واتجاهاته الموضوعية أو امتداداته النوعية داخل مواد معلومات الويب. فالبحث بالمعنى والمقابل والجذر يجعل المفردة الواحدة في استراتيجية البحث مجموعة مفردات بحثية في آن واحد، ومن ثم فإن نظام الاسترجاع يصعب عليه تتبع مفردة البحث في شتاتها الشكلي واللفظي بما يحدد من قدرة النظام على التحكم بمفردة البحث وتتبعها منعا أو تمريرا، ومن ثم فإن نتاج هذه الصعوبة هو انخفاض مستوى التدخل والتحيز من نظام الاسترجاع في التأثير على النتائج ومواد المعلومات المسترجعة. وتأتي الصعوبة في تحقيق التحيز لتوجهات النظام من شتات المشتقات ونواتج عملية التحليل اللغوية لمفردات البحث، خاصة تلك التي تنتمي الى لغة ثرية بالمشتقات قوية البناء اللغوي مثل اللغة العربية، وعلى ذلك فتوزيع البحث على مفردات غير مباشرة لم تدخل في استراتيجية البحث ككلمات مفتاحية ثم التطابق بينها وبينكافة كلمات المحتوى النصي لمادة المعلومات حيث

التكشيف الاشتقاقي، يجعل من الصعوبة بما كان تتبع النظام لهذه المصطلحات والتحكم في عملية البحث والاسترجاع ومن ثم التحكم في النتائج المسترجعة[51 52].

ثالثا : البحث بالويب الدلالية : ترتكز آليات البحث بالويب الدلالية على استخدام المفاهيم والدلالات الموضوعية لمفردات البحث بدلا من استخدام مبدأي المضاهاة والتحليل اللغوي، فكلا المبدأين في البحث يدور حول المفردة ذاتها والاشتقاقات التي تنبثق منها، أما الويب الدلالية فتستخدم المفردة كمرتكز ونقطة انطلاق لتلون البحث بلون الموضوع دون الوقوف على المفردات التي تحمله. وتستخدم الويب الدلالية مجموعة أدوات برمجية هي **XML. RDF, OWL** اللغة المعيارية الممتدة وإطار وصف المصادر ولغة انطولوجيا الويب، وهذه جميعها تتكامل من خلال أداء كل منها لمهام محددة في كون الأولى النظام الترميزي لكتابة البيانات ونشرها على الويب، أما الثانية فهي تضبط العلاقات بين المصطلحات وتحدد هوية المفردات ودلالاتها في شبكة مفاهيمية أما لغة انطولوجيا الويب فالتي ترسم طبيعة البنية الدلالية المتقدمة لتفعيل الويب الدلالي. وعلى أثر ذلك فنظام الاسترجاع يبقى عاجزا عن تتبع مختلف المفردات المرتبطة دلاليا لكونها ترتبط من حيث المفهوم والموضوع وليس من خلال الانبثاق الشكلي أو اللغوي للكلمات، وبالتالي يظل تحيز نظام الاسترجاع وتوجهاته حبيسة الألفاظ والمفردات اللغوية التي اعتمدت فيما سبق على تبني مبدأ المضاهاة والتطابق الشكلي لتتبع الموضوع بتتبع مفرداته. فما يزيد أمر التحيز تعقيدا في نظم الويب الدلالي هو قاعدة المعرفة **KB** التي تمثل داعما متقدما مع الكشاف المقلوب في قاعدة بيانات النظام، بحيث يمدها الكشاف بمختلف المفردات التي تنتج عن تحليل مواد المعلومات، ثم يستقبل منها الارتباطات الدلالية لمفردات البحث، ومن ثم فإن قدرة النظام على تعمد التحيز هنا يمثل صعوبة بالغة لوجود أكثر من جسر لتحقيق التباديل والتغيير بين مفردات قاعدة المعرفة داخل النظام[53].

5 / 2 / 2 الفرز والترتيب ranking algorithm

تكتمل آليات البحث والاسترجاع بالوقوف عند جودة الفرز والترتيب لنتائج البحث المسترجعة على واجهة النظام ؛ حيث تضطلع عملية الفرز بانتقاء مجموعات مواد المعلومات التي تتوافق واستراتيجية البحث وحاجة المستخدم دون غيرها، بل وانتقاء مجموعات المواد المتوافقة موضوعيا واستفسار البحث دون مواد المعلومات التجارية أو الدعائية، كما أن عملية الترتيب تختص بتقديم وتأخير مواد المعلومات المسترجعة وفقا لأولوية القيمة والأهمية الموضوعية في تلبية الحاجات الموضوعية. وقد استقر مفهوم الفرز او التصنيف كما يطلق عليه **Ranking** عند كونه قدرة نظام الاسترجاع على جمع وانتقاء النتائج قوية التي تتصل موضوعيا باستفسار البحث، التي تكفي الحاجات المعلوماتية لمستخدمي النظام. ومن ثم فكفاءة النظام تتمثل في أداء الفرز والترتيب بشكل امثل على وجهين هما أولا انتقاء مواد المعلومات تتصل مباشرة بموضوع البحث بما يفيد في تغطية الحاجات المعلوماتية، ثانيا استبعاد مختلف المواد التي قد تؤثر سلبا على قيمة النتائج المسترجعة أو التشويش على المستخدم وتصيبه بالتشتت. وعلى ذلك يمكن اعتبار الفرز مقياس كفاءة النظام في

الاسترجاع وتلبية الحاجات المعلوماتية، كما أنه مؤشر دقيق لقياس درجة تحيز النظام موضوعيا وقياس مدى التدخل في تمرير أو منع نتائج بذاتها وفقا لرغبة النظام في ذلك. ولعل شأن الفرز شأن باقي الإجراءات السابقة في نظام الاسترجاع تتأثر بكيفية الاداء ونوعية برمجيات العمل المستخدمة في تكشيف أو بحث أو فرز مواد المعلومات ؛ حيث تقوى قدرة التحيز في النظام كلما استخدم مبدأ المضاهاة والمطابقة وفقا للألفاظ والمفردات، أما حال العمل وفقا لتقنيات التحليل الدلالي فإن قدرة التحيز داخل النظام تدنو إلى أقل درجاتها[54].

دعمت برمجيات الويب الدلالي آليات الفرز في نظم الاسترجاع على لتطوير آليات الفرز من تلك المعتمدة على الإطار الشكلي اللفظي المصطلحي لكلمات البحث والكلمات الدالة في مواد المعلومات إلى الشكل المفاهيمي الموضوعي الدلالي الهادف لتتبع الموضوع والاحتياجات المعلومات الممثلة في استراتيجية البحث والكامنة في الكلمات الدالة الكشفية لمواد المعلومات داخل قاعدة بيانات النظام. وبذلك تتحول قدرة النظام من كفاءة تتبع المصطلحات إلى ضعف تتبع الموضوع وتوجهاته الدلالية التي تمثلها شبكة المفاهيم والمفردات الدالة في الويب الدلالي بما يعود على النظام بفقد السيطرة على النتائج المسترجعة والتحكم في طرائق عرضها وتشكيل نتائج البحث وفقا لأولوياته المتحيزة وتوجهاته الفكرية. وقد تعددت آليات تفعيل **الشكل الأول للفرز المصطلحي الكلمي** في تتبع تردد المصطلح والمناطق الموضوعية ووزن أهمية المصطلح في استفسار البحث. وتتفق جميع هذه الآليات في تقدير أهمية مصطلحات البحث ووزن أهمية مادة المعلومات في الرد على الاحتياجات الموضوعية وفقا للمصطلح والمفردات ذاتها، وفقا للآليات التالية :[55-56]

1) تردد المصطلح **term frequency** تعمل آلية تردد المصطلح على وزن أهمية مادة المعلومات وفقا لمرات تكرار وتردد مصطلحات البحث في كل مادة بشكل مستقل، بحيث تكون المواد الأكثر احصاءً في تردد مصطلحات البحث هي المواد الأقرب صلة بموضوع استفسار البحث.
2) موقع المصطلح **Zone weighting** تستخد النظم آلية تحديد أهمية مادة المعلومات وفقا لموقع تواجد مصطلحات البحث داخل مادة المعلومات، حيث تختلف درجات الأهمية لمواقع مادة المعلومات فيأتي في أعلاها عنوان المادة المعلومات ثم موقع الكلمات المفتاحية ومواقع كتابة المتن ومجال او اسم المادة. **Anchor Text, Title Text, Document's URL, Headings , Page Description**
3) درجات التقارب للمصطلحات **Term proximity scoring**: تهتم آلية الفرز هذه بدرجة التقارب في موقع مصطلحات البحث داخل متن مادة المعلومات ؛ حيث لا يكفي مرات ترددها أو مواقعها المشتتة في وزن أهمية مادة المعلومات، إنما يجب أن تكون هذه المصطلحات أقرب ما تكون من حيث مواقع ذكرها بشكل متقارب ليعزز أكثر هوية السياق والموضوع التي تشكله معا ومجتمعة في جزء واحد من مادة المعلومات.

4) الجمع بين الموقع والتقارب **Combining term proximity with zone weighting** تجمع هذه الآلية بين جانبي تقارب مصطلحات البحث في مادة المعلومات ومواقع تواجد هذه المصطلحات في مادة المعلومات بحيث يتكامل موقع ذكرها ودرجة تقاربها في وزن أهمية مادة المعلومات ومن ثم تحديد درجة الصلة الأقرب لموضوع البحث، بل وتحديد قيمة ترتيبها في نتائج البحث على واجهة النظام.

يرتكز **الشكل المفاهيمي الدلالي** في آليات الفرز على أساس الاسترجاع بمحتوى وتوجه استفسار البحث، وهو شكل يتجاوز الإطار اللفظي او المصطلحي لمفردات البحث ليصل إلى مستوى تحليل المفاهيم والسياق الموضوعي لمفردات البحث. فتمثل استراتيجية البحث سياقا وتوجها موضوعيا لمستخدم النظام حيث يعكس ترابط كلمات البحث وسياق صياغتها ضمنيا توجها بذاته أعمق في الدلالة من اختيار المفردات أو مجرد تجميعها معا في استفسار بحث واحد، وقد تطور أداء الفرز ليعمل على انتقاء مواد المعلومات وفرزها وفقا لمحتواها وتوافقها ليس مع الكلمات وإنما مع التوجه الموضوعي والمحتوى الكامن في سياق استفسار البحث[57]. ويعزز هذا الشكل الدلالي حيادية النظام في تقييم وفرز وترتيب نتائج البحث المسترجعة بما يتوافق واحتياجات المستخدم وتوجهاته الفكرية دون أن يترك للنظام فرص التحيز والتدخل في فرز وانتقاء مواد المعلومات. ولقد بنت نظم الويب توجهها الدلالي الموضوعي في الفرز على ثلاثة دوافع أساسية تتمثل في ؛ الأول التوجه نحو الموضوع والفكرة التي يتضمنها استفسار البحث في مجموع مفرداته. الثاني الأخذ بأدوات التحليل اللغوي والمعجمي للتعامل مع مفردات استفسار البحث لاستخلاص الدلالة اللغوية منها، الثالث التأكيد على التوجه الموضوعي الأكثر وضوحا في مجال استفسار البحث وفقا لوزن المفاهيم داخل مفردات البحث وقياس أوزان التوجهات الفكرية التي يستهدفها الاستفسار ثم العمل على قياس مواد المعلومات اعتمادا عليه. وقد عنيت دراسة **Leonidas Akritidis** بطرح آلية **Query Independent Quality Scores (QUIQS)** لوزن قيمة مفردات البحث وتطبيق آليات الفرز اعتمادا عليها، ويتم ذلك من خلال مقومات ثلاث هي :[58]

1. قيمة مادة المعلومات **material value** حيث تختلف مواد معلومات موضوع ما في قيمتها وفقا لوزن أهمية كل مادة وقيمتها في تلبية احتياجات مستخدم النظام، وتأتي أهمي كل مادة من حيثيات مثل كم الاستشهادات بهذه المادة في هذا الموضوع، أيضا حجم مشاركة مستخدمي المادة إذا ما كان المحتوى يسمه بإبداء وجهات نظرها خاصة في المحتوى التفاعلي، أيضا المصادر التي اعتمدت عليها هذه المادة.
2. قيمة المنتج **author influence** حيث تختلف قيمة مادة معلومات وأخرى بحسب قيمة المحرر أو منشئ مادة المعلومات فتعلو قيمة مادة على اخرى بحسب قيمة كاتبها بما يجعلها أفضل في تلبية احتياجات المستخدم وومن ثم تدعمها آليات الفرز.

3. معامل تأثير النشر **publishing impact** حيث يعكس موقع نشر مادة المعلومات أو الجهة الناشرة لمادة المعلومات قيمة محتوى هذه المادة ومدى الثقة في محتوى المادة عن غيرها من مواد المعلومات التي تحمل محتوى غير موثوق به أو يشوبه ضعف نتائجه ومضمونه.

المبحث الثالث التطبيقي : قياس التحيز في نظم استرجاع الويب:

يهدف هذا المبحث لدراسة واقع ظاهرة التحيز في نظم استرجاع الويب وبالتحديد في فئة محركات البحث دون غيرها،ذلك لكونها الأكثر ارتباطا وحيثيات الدراسة في برامجها الآلية وإجراءات العمل. وقد تمت هذه الدراسة في أعقاب الانتخابات الامريكية الأخيرة عام 2016 في الفترة بين ديسمبر 2016 إلى مارس 2017، اعتمادا على عينة من محركات البحث بلغت ستة محركات كانت الأعلى في ترتيب محركات البحث عالميا وفقا لموقعي التصنيف العالميين **Alexa Global Traffic Rank, and U.S. Traffic Rank** التي أصدرت تصنيفها في يناير 2017 على **www.alexa.com**، وقد اقتصرت العينة على ستة فقط من محركات البحث لكونها تحمل اعلى معدل زيارات شهري لمستخدمي الويب بلغت في أعلاها مليار وثلاثمائة مليون لمحرك **Google** وأقلها مائة مليون زائر لمحرك البحث **wow**.[59] وعلى ذلك فستكون عينة الدراسة من نظم استرجاع الويب هي محركات البحث الأكثير استخداما وفقا لموضع أليكسا، ممثل في الجدول التالي :-

جدول (2) عينة الدراسة من نظم استرجاع الويب وفقا لمدلات الاستخدام لتفييم alexa

م	نظام الاسترجاع	الرابط	معدل الزيارة شهريا
1	GOOGLE	/https://www.google.com.eg	1,600,000,000
2	BING	/http://www.bing.com	400,000,000
3	YAHOO	/https://search.yahoo.com	300,000,000
4	ASK	/http://www.ask.com	245,000,000
5	AOL	http://search.aol.com/aol/webhome	125,000,000
6	WOW	/http://www.wow.com	100,000,000

وقد اعتمدت الدراسة التطبيقية على استخدام أداتي الملاحظة والمعايشة لمختلف عمليات البحث التي اجريت على عينة الدراسة والتي اجريت على كل محرك لاختبار مختلف عناصر الدراسة التطبيقية. وعملت الدراسة التطبيقية على استحضار مختلف جوانب ظاهرة التحيز وتمثلت في محاور الدراسة التجريبية، وقد تنوعت بين تلك التي تقيس اتجاهات نظم الاسترجاع من حيث قابلية التحيز ودرجة تأثيره أو انتفاؤه داخل النظام، أيضا عناصر لقياس مستوى التحيز في كل مرحلة من مراحل واجراءات العمل في برامج النظام، ثم عناصر لقياس الشكل النهائي المنعكس على النتائج المسترجعة على واجهة البحث والاسترجاع التي يمكن لمستخدم النظام تحديدها ومراقبتها. ولتحقيق أهداف الدراسة التطبيقية فقد ركزت أكثر محاورها على دراسة وتحليل آليات البحث وفحص مجموعات النتائج المسترجعة على كل نظام وفقا لطبيعة المحور موضع التطبيق والدراسة. وسوف تعتمد الدراسة على استقراء اجراءات النظام وتتبع ملامح التحيز فيها لكونها تتم داخليا في كل النظم

ببرمجيات وتقنيات غير منظورة، أما تتبع ملامح التحيز وأثره فسيكون من خلال دراسة النتائج واتباع معادلات قياس التحيز وفقا للمعادلات والمعطيات الاحصائية الرقمية للنتائج المسترجعة.

3 / 1 قابلية التحيز في النظام :

تعكس بعض الملامح في نظام الاسترجاع موقف النظام من التحيز من حيث المبدأ، ومدى قابلية النظام لاصباغ إجراءاته وعملياته بصبغة التحيز. فثمة ملامح مثل وجود الاعلانات الترويجية أو الدعائية للخدمات والسلع على واجهة النظام، تجعل النظام مقيدا بتقديم المصلحة التجارية والخدمات المرسمة على حيادية التعامل مع مواد معلومات الويب خاصة إذا ما كانت تقع في فلك التوجهات التجارية والتسويقية لمواد المعلومات الدعائية. ويتضح من واجهة نظام الاسترجاع تجلي بعض المصطلحات التي تعلن عن هوية النظام وطبيعته المتحيزة سواء لقضايا أو توجهات أو لمصالحه التجارية والإعلانية، ومن ثم فوجود مصطلحات مثل ترويجي أو مدفوع الأجر تعني تحيز النظام لإعلاء هذه المادة والمعلومات على غيرها من مواد المعلومات الأخرى. كما أن بعض البوابات الخدمية المقدمة من جانب النظام كخدمات إضافية مجانية لمستخدميه، كخدمات المدونات على نظام الاسترجاع **google** وقنوات التواصل الاجتماعي مثل **Igoogle** تجعل نظام الاسترجاع متحيزا لتكشيف ونشر هذه المواد حتى يتم ترويجها ورفع معدل زياراتها ومن ثم مضاعفة قيمة الإعلانات التي تنشر على هذه الصفحات. أما آخر الملامح التي تعلي من قابلية التحيز داخل النظام فتتمثل في تابعية النظام وانتمائه المؤسسي، فوفقا لطبيعة الشركة والهيئة الداعمة والقائمة على النظام تكون توجهاته في التعامل مع مواد المعلومات سواء أكانت هذه التوجهات إقتصادية أم سياسية أم أيديولوجية. ويوضح جدول (2) مجموعة الملامح التي إذا ما برزت في نظام الاسترجاع دعمت قابلية النظام لتفعيل التحيز في معالجة وبحث مواد معلومات الويب.

جدول (3) قابلية التحيز في النظام في عينة الدراسة

قابلية التحيز في النظام		نظم استرجاع الويب					
		GOOGLE	BING	YAHOO	ASK	AOL	WOW
توجهات التحيز	معلنة	√	---	√	√	√	---
	غير معلنة	---	√	---	---	---	√
الاعلانات		√	---	√	√	√	---
كلمات الترويج		√	---	√	√	√	---
تابعية النظام		---	---	---	---	---	---
البوابات الخدمية		√	√	√	---	√	√

يتضح من جدول (3) بعض المشاهدات مثل :-

أولا : تضمنت أربعة نظم من عينة الدراسة ممثلة أغلب العينة ملامح قابليتها للتحيز، وبخاصة التحيز التجاري الهادف لترويج مواد معلومات إعلانية. وتمثل التحيز التجاري في دعم المواد المعلنة والتنويه عن ترويجها تحت مسمى إعلانات باستخدام مصطلح (مصر للطيران)، وقد دلل على التحيز استباق

هذه النتائج في الترتيب لباقي نتائج البحث بما يمكن أن يهيئ للمستخدم أنها أعلى النتائج المسترجعة ومن ثم الاهتمام بالدخول عليها عن غيرها.

ثانيا : أصبغت البنية العامة لنظم الاسترجاع وتحولها إلى بوابات متكاملة الخدمات طبيعة النظام بتحوله من نظام استرجاع مخصص لبحث مواد الويب إلى بوابة خدمات ينتج عن خدماتها مواد معلومات مثل صفحات المدونات والتواصل الاجتماعي، بما يجعل حرص النظام على تقديم هذه المواد كأولوية في الاسترجاع عن باقي مواد الويب.

ثالثا : أثبتت ملاحظات القابلية للتحيز أن نظم الاسترجاع لا تكتفي بدورها في معالجة وبحث مواد الويب، وإنما شملت أيضا التعاطي مع مستخدميها ومواقع التواصل ومواقع النشر الإعلامي، بما جعلها نظما منفتحة تتخاطفها توجهات وتحيزات نوعية تجارية بهدف الترويج أو أيدولوجية بواقع التداخل مع مواد الويب، وهو ما يشير إلى قابليتها الأكثر للتحيز في بحث واسترجاع النتائج.

3 / 2 آليات الإضافة والانتقاء :

يبدأ تفعيل التحيز في نظام الاسترجاع مع توجهات النظام نحو اختيار متعمد لمواد معلومات الويب، ويرتبط بالاختيار المتعمد الامتناع أيضا المتعمد في إضافة مواد معلومات بذاتها. ونظرا لكون نظام استرجاع الدراسة تنطلق في عملها من برامج وأدوات تقنية تمارس عملها من منطلق التعليمات والتحرك البرمجي وليس منطلق الاختيار والانتقاء، فقد ادخلت نظم استرجاع الويب جوانب تطوير وخصائص برمجية تجعل نظم الاضافة كبرامج الزاحف تعمل على التحرك العنقودي في مجال موضوعي محدد بكل فروعه، وليس التحرك الافقي بين مواقع الويب وفقا لما بينها من روابط تقنية **Links**. وعلى ذلك فطبيعة التخصص الموضوعي أو تخصص مواد المعلومات داخل نظام الاسترجاع تعكس بدورها توجهات النظام في اختيار واضافة مواد المعلومات مسبقا، وعلى أثرها سيكون طبيعة النتائج المسترجعة من جنس محتوى قاعدة بيانات النظام، وهو بذلك قد فرض مسبقا توجهاته على مستخدميه مع بداية إضافة مواد المعلومات وبناء قاعدة بياناته. ومن أهم الجوانب التي تعكس تحيز النظام في الإضافة والاختيار نوعية مواد المعلومات مثل الإعلامية لخدمة توجهات محددة أو دعائية للترويج لمنتجات أو سلع بمقابل للنظام، أو أن يتحيز النظام لمواد المعلومات الخاصة بخدمات النشر والتدوين المقدمة على واجهة النظام ذاته.

جدول (4) توجهات وآليات الإضافة والانتقاء في عينة الدراسة

آليات الإضافة والانتقاء		نظم استرجاع الويب					
		GOOGLE	BING	YAHOO	ASK	AOL	WOW
التغطية والشمول	متخصص	---	---	---	---	---	---
	عام	√	√	√	√	√	√
مواد المعلومات الشخصية		√	√	√	√	√	√
	علمية	√	√	√	√	√	√

توعية المواد	إعلامية	√	√	√	---	√	√
	دعائية	√	---	√	√	√	---
المواد الإعلامية	متضمنة	√	---	√	---	√	√
	منفصلة	√	√	---	---	√	√
مواد معلومات النظام		√	---	---	---	---	---

يعكس جدول (4) بعض الجوانب المهمة وهي :-

أولا : أظهرت طبيعة النتائج المسترجعة من عينة الدراسة توجه نظم الاسترجاع نحو الشمولية والتغطية العامة للمجالات الموضوعية، وهو ما انعكس على شمولية مواد الويب النوعية موضوعا وشكلا. وقد ضمت نظم العينة بين مجموعاتها صفحات المعلومات الشخصية والتواصل الاجتماعي لبعض أهم مستخدميه، ويشير هذا الجانب إلى قدرات أكبر من التحكم في النتائج المسترجعة وتقديم نتائج على أخرى، بما يعزي بوجود تحيز الإضافة والانتقاء بين المواد التي يضيفها النظام ثم يقوم على اتاحتها لمستخدميه مرة أخرى.

ثانيا : تنوعت مواد المعلومات بين العلمية والإعلامية والدعائية، وقد تجلى ذلك في جميع نظم العينة بما يعكس قدرة النظام على التخير في الاسترجاع وتغليب نوع على آخر من مواد المعلومات النوعية وفقا لتوجه النظام وطبيعة محتوى مادة المعلومات المسترجعة. وقد بدا التحيز لمواد النظام التي تنشا وتنشر على موقعه بين خدماته أهم النتائج المتحيز لها في الاسترجاع، كما ظهر في حالة نظام **Google** باسترجاع صفحات من مدوناته المتاحة على نظامه.

ثالثا : تمثل المواد الإعلامية أكثر مواد الويب حملا للمحتوى المتحيز ؛ حيث تتسم الأخبار والمواد الإعلامية بالعرض الأحادي للموضوعات والأحداث، ومن ثم فهي مادة ثرية بالمعلومات التي توافق توجهات أو تحيز نظام الاسترجاع. ولذا حرصت أغلب نظم العينة على إتاحة البحث في المواد الإعلامية بشكل مستقل عن مواد الويب الباقية مع تضمين أبرز المواد الإعلامية ضمن نتائج البحث.

3 / 3 إجراءات التكشيف والمعالجة:

تلعب إجراءات التكشيف والمعالجة الفنية الدور الأساسي لتفعيل التحيز في نظام الاسترجاع إذا ما أراد النظام فرض توجهاته وتحيزه على عمليات البحث والاسترجاع أمام مستخدمي النظام. وبرغم التحكم الرقمي من برامج المكشف **Indexer** في اختيار وانتقاء الكلمات الدالة على هوية ومحتوى مادة المعلومات، غير أن التقنيات الحديثة في عمليات التكشيف تمكن النظام من تعديل وجهات التكشيف والتحكم المسبق في طبيعة الكلمات الدالة أو المصطلحات الكشفية التي تعبر عن محتوى الوثيقة، والتي ستكون وحدها فيما بعد الوسيلة الوحيدة لاسترجاع واستدعاء هذه المواد ضمن نتائج البحث. وعلى ذلك فإن تحديد آلية التكشيف وانتقاء المفردات الدالة هو يد النظام الطولى في فرض تحيزه وتوجهاته على تنظيم وبحث مواد المعلومات، ومن ثم فإن النظام يحدد بتحيزه الوجهة الموضوعية التي ستخدمها مادة المعلومات، بل وتحديد استراتيجية البحث التي يمكن لمادة المعلومات

دخولها ضمن نتائج البحث وفقا لمفردات بحثية محددة. ورغم كون عملية التنظيم والتكشيف تتم داخل نظام الاسترجاع ووفقا لاجراءاته الرقمية، غير أن ثمة شواهد يمكن استخدامها لقياس ورصد التحيز في أجراءات التكشيف لمواد الويب، وأهم هذه الشواهد ما يتمثل في طبيعة المفردات الكشفية المخصصة للتعبير عن مواد الويب، إذا ما كانت تتمثل في الكلمات أم المفاهيم ؛ فالكلمات تعبر عن قدرة التحيز على التحكم في معالجة مادة المعلومات ومن ثم الشك في بسط سيطرة توجهات النظام على مواد المعلومات، أما المفاهيم فتعكس تقدما في معالجة مادة المعلومات بما يجعلها خارج سيطرة النظام على التحكم في محتوى مواد المعلومات. أيضا نمط التكشيف النوعي المتبع قد يسهل معه أو يعصب تفعيل التحيز على معالجة مواد الويب، حيث يعني التكشيف الاحصائي فرصة أكبر للنظام باستخدام قوائم المرور والمنع في اختيار الكلمات الكشفية الدالة على محتوى مادة المعلومات، ومن ثم التحيز ضد او مع توجه فكري محدد متمثلا في المصطلحات أو الكلمات الدالة عليه، ويخالف ذلك التكشيف الدلالي الذي لا يرتبط بمفردات محددة بل يسمح لمادة المعلومات بالمرور ضمن نتائج البحث أو التوقف عن ذلك وفقا لطبيعة محتواها وليس تبعا للائحة المنع أو التمرير للنظام ذاته. وينطبق الأمر ذاته على تكشيف مجموعات الوسائط المتعددة ؛ حيث مازالت تعتمد نظم الاسترجاع على تكشيف الوسائط المتعددة تبعا للنصوص المصاحبة في صفحة أو موقع الويب التابعة له هذه المواد.

جدول (5) تأثير إجراءات التكشيف والمعالجة في عينة الدراسة

إجراءات التكشيف والمعالجة		نظم استرجاع الويب					
		GOOGLE	BING	YAHOO	ASK	AOL	WOW
مستوى التكشيف	كلمات	√	√	√	√	√	√
	مفاهيم	---	---	---	---	---	---
أدوات التكشيف	المحللات	---	---	---	---	---	---
	المعاجم	√	---	√	---	---	---
نمط التكشيف	احصائي	√	√	√	---	√	√
	تحليلي	---	---	---	---	---	---
	دلالي	---	---	---	---	---	---
وجهات التكشيف	المحتوى	√	√	√	---	---	√
	المحارف	√	√	√	√	√	√
تكشيف الكائنات	الصور	√	√	√	---	√	√
	الفيديو	√	√	√	---	√	√
	الصوت	√	---	---	---	---	---

يتضح من جدول (5) ما يلي :-

أولا : اعتمدت عينة الدراسة جميعها على التكشيف بمستوى الكلمات دون المفاهيم، بما يعني منح النظام فرصة أكبر للتحكم في كلمات ومصطلحات بعينها ومن ثم التحيز بالتمرير أو بالمنع لبعض

المفردات الكشفية لمواد الويب. ولم يستخدم أي من نظم العينة مستحدثات الويب الدلالي أو لغة الانطولوجيا في التكشيف، كونها تبعد النظام كثيرا عن التحكم في تكشيف مواد الويب ومن ثم خروج نتائج البحث عن سيطرته.

ثانيا : لم تستخدم أي من نظم العينة أدوات المعالجة اللغوية من محللات ومعاجم إلا نظامي **google, yahoo** في معالجة وتهذيب مفردات البحث ؛ ويرجع ذلك للعنصر السابق كون نظم الاسترجاع مازلت تعتمد الكلمات بألفاظها وشكلها أكثر من دلالتها ومفاهيمها في عملية التكشيف والاسترجاع لسهولة حجبها وتمريرها، وتطبيق توجهات النظام على النتائج المسترجعة.

ثالثا : ارتباطا بالتكشيف على مستوى الكلمات فقد التزمت عينة الدراسة بنمط التكشيف الاحصائي القائم على احتساب قيمة الدلالة والأهمية الموضوعية للكلمات وفقا لعدد مرات ترددها وتكرارها في متن مادة المعلومات، ومن ثم يمكن للنظام تحديد بعضا من الكلمات التي توافق توجهات النظام لاختيارها كلمات كشفية بديلا عن الكلمات الأعلى ترددا في حالة عدم موافقتها لتوجهات النظام.

رابعا : اعتمدت عينة الدراسة على استخدام محارف العنوان والكلمات المفتاحية لمادة المعلومات في إجراء التكشيف أكثر من اعتمادها على المحتوى ذاته، وقظ انعكس ذلك من إظهار كلمات البحث في النتائج المسترجعة التي قد وردت في متن مادة المعلومات. وقد بدا ذلك واضحا في نظامي **ASK, AOL**.

خامسا : اهتم نظام **google** وحده بتكشيف كائنات الوسائط المتعددة المتضمنة في متن مواد الويب، ورغم كون أغلب نظم العينة تبحث في الصور والفيديو إلا أن تكشيف هذه الوسائط وبحثها يعتمد على متن مادة الويب ونصوصها وليس الميتاداتا المتضمنة مع هذه الوسائط. ويمثل تكشيف هذه الوسائط عاملا إيجابيا في دعم حيادية النظام ؛ حيث تؤثر ميتاداتا هذه الوسائط في تحديد الصلة الموضوعية لمادة الويب بعيدا عن تحيز النظام. وقد أظهرت نتائج عينة الدراسة افتقاد نظم الاسترجاع لتكشيف الوسائط المتعددة وقصر الكلمات الدالة على نصوص مواد الويب بما يعني قدرة أكبر للنظام في التحكم والتحيز معها.

3 / 4 تقديم المقترحات والاكمال التلقائي:

تمثل نتائج البحث انعكاسا لطبيعة المفردات ونتيجة اجمالية لصياغة استراتيجية البحث المدخلة للنظام، حيث يلتزم النظام باستدعاء النتائج التي توافق فقط نطاق وملامح الكلمات المفتاحية سواء في طبيعة الكلمات ودلالاتها أو في ترتيب الكلمات في سياق استفسار البحث. وعلى ذلك فإن النتائج المسترجعة لمستخدم النظام لا تبدأ عند عمليات الفرز والترتيب الداخلية، بل تبدأ عند إختيار وإدخال استفسار البحث والكلمات المفتاحية إلى النظام، ومن ثم فإن النظم المتحيزة ذات التوجهات الخاصة تعمل على التدخل في اختيار كلمات البحث وتشكيل استفسار البحث بما ينعكس على طبيعة وقيمة النتائج المسترجعة لمستخدم النظام. وقد اعتمدت نظم استرجاع الويب في تفعيل هذا التدخل والتحكم في كلمات البحث على آليات عمل متطورة تمثلت في آلية الإكمال التلقائي **autocomplete** التي تختص بتقديم كلمات مقترحة وسياقات بديلة أمام المستخدم بمكنه الاختيار

من بينها لإدخال استفسار البحث. ويستهدف النظام المحايد من هذه الآلية مساعدة ودعم المستفيد في اختيار وانتقاء الكلمات المفتاحية الأقرب للتعبير عن احتياجاته المعلوماتية، غير أن النظم ذات التوجهات والتحيز تستخدم هذه الآلية لتعمد إقحام المفردات التي تخدم توجه وتحيز النظام، وعلى أثر ذلك يتم التحكم مسبقا في طبيعة وهوية النتائج المسترجعة للمستخدم. ويبقى العمل بهذه الآلية مقيدا بجوانب مختلفة يتحكم بها النظام لاختيار المفردات مثل عمليات البحث السابقة والنطاق الجغرافي الذي يخدمه النظام، غير أن استخدام مقارانات العمل بهذه الآلية بين النظم وبعذها البعض قد اعتمدت عليه الدراسة الحالية في طرح مؤشر لقياس درجة التحيز في نظام الاسترجاع من خلال الربط بين عدد المقترحات التي يتميز بها النظام وتعبر عن توجهاته وتعرف بالمقترحات النوعية، بالنسبة لإجمالي عدد المقترحات المقدمة من النظام. وقد مثلت المعادلة التالية قياس مؤشر تحيز نظم الاسترجاع في تقديم المقترحات لعينة الدراسة :-

مؤشر التحيز بالمقترحات = عدد المقترحات النوعية ÷ عدد المقترحات المقدمة

جدول (6) تقديم المقترحات والاكمال التلقائي في عينة الدراسة

تقديم المقترحات للبحث		نظم استرجاع الويب					
		GOOGLE	BING	YAHOO	ASK	AOL	WOW
islam religion	العدد الكلي	10	---	10	8	8	0
	عدد المختلف	2	---	1	1	3	0
	نوعية التحيز	سلبي	---	سلبي	سلبي	محايد	0
	النسبة	%20	---	%10	%13	%38	0
president trump	العدد الكلي	10	---	10	0	8	2
	عدد المختلف	2	---	5	0	1	1
	نوعية التحيز	سلبي	---	محايد	0	محايد	محايد
	النسبة	%20	---	%50	0	%13	%50
Syria	العدد الكلي	10	---	10	8	8	8
	عدد المختلف	1	---	6	3	1	1
	نوعية التحيز	سلبي	---	محايد	محايد	محايد	محايد
	النسبة	%10	---	%60	%38	%13	%13
palestine israel	العدد الكلي	10	---	10	4	8	0
	عدد المختلف	1	---	6	0	0	0
	نوعية التحيز	سلبي	---	ايجابي	0	0	0
	النسبة	%10	---	%60	0	0	0
	العدد الكلي	10	---	10	2	8	0

Arab Spring	عدد المختلف	0	---	1	0	1	0
	نوعية التحيز	0	---	محايد	0	محايد	0
	النسبة	0	---	%10	0	%13	0
terrorism	العدد الكلي	10	---	10	8	8	8
	عدد المختلف	1	---	6	1	3	5
	نوعية التحيز	سلبي	---	محايد	سلبي	محايد	سلبي
	النسبة	%10	---	%60	%13	%38	%63

بالاعتماد على بيانات الجدول (6) يتضح الآتي :-

أولا : تباينت نظم استرجاع علينة الدراسة بشكل جلي في التعامل مع آلية الاكمال التلقائي والتحيز في توجيه مستخدم النظام ؛ حيث كان التباين نوعيا في إيجابية وسلبية التحيز لبعض القضايا، وأيضا كميا في عدد المقترحات المقدمة من كل نظام. وقد كان نظام **google** أعلى نظم عينة الدراسة تحيزا نوعا وكما، حيث غلبت على **google** الميل السلبي في التعامل مع بعض القضايا التي تخص العالم العربي وإن كان عدد المقترحات السلبية ليس كثيرا في حدود مقترح أو اثنين فقط. بينما لم يستخدم تماما نظام الاسترجاع **BING** آلية الإكمال التلقائي في أجراء عمليات البحث.

ثانيا : شكلت نسب الاختلاف في نظام **yahoo** أعلى نسب الاختلاف النوعي والعددي في المقترحات المقدمة مقارنة مع **google**، ففي الوقت الذي قدم فيه **google** مقترحات سلبية مثل اقتراح **islam** في مقترحات **terrorism**، أيضا اقتراح **Simpsons** الشخصية الكرتونية عند ادخال اسم الرئيس الأمريكي **trump**، قدم نظام **yahoo** مقترحات إيجابية مثل **pease treaty** عند إدخال **palastine**.

ثالثا : أثبتت مشاهدات آلية تقديم مقترحات البحث في نظم العينة أن نظم الاسترجاع تستطيع التحكم تماما في نوعية وطبيعة المقترحات المقدمة لمستخدم النظام سواء أكانت مقترحات سلبية أم إيجابية لخدمة قضية البحث. وذلك بدا جليا من الاختلاف النوعي التام بين نظامي **google** و**yahoo** و باقي النظم الأخرى. ويمثل ذلك دلاليا على تفعيل مبدأ التحيز في نظام الاسترجاع بمقترحات سلبية أو إيجابية عند التعامل مع مفردات بحث محددة تمثل قضايا يراها النظام بوجهته ورؤيته الخاصة. وقد استخدمت الدراسة وصف سلبي للمقترحات التي تمثل جنوحا تاما عن حيادية الموضوع، أما وصف محايد فيستخدم في حالة الاختلاف الموضوعي فقط دون تحيز سلبي أو إيجابي.

رابعا : لم تقدم نظم **WOW, AOL, ASK** المقترحات النوعية والعددية بالقدر المؤثر كنظم استرجاع لمواد الويب، وهو ما يؤكد حقيقة تراجع ترتيب هذه النظم في التصنيف العالمي لمحركات البحث. بالاضافة إلى أن آلية المقترحات في هذه النظم لم تعمل تقنيا بالشكل المرجو فجميعها اعتمد على استكمال السياق وفقا لحروف الادخال لمفردات البحث، وليس اقتراحا للسياق المكون من مفردات تمثل وجهات تناول موضوع البحث. ولذلك وصلت مقترحات هذه النظم لمستوى الصفر في التعامل

مع بعض المفردات مثل **Arab Spring** و **Israel palestine** فلم تقدم لهذه المفردات أية مقترحات سلبية أو إيجابية.

خامسا : أظهر ثبات أعداد ونسب التحيز في نظم عينة الدراسة أن آلية تقديم المقترحات تعمل بشكل ممنهج وغير عشوائي، وتأكد ذلك أن غالبية النسب والاعداد كانت ثابتة برغم اختلاف مفردات البحث المستخدمة التي تعبر في كل مرة عن قضية وموضوع مختلف تماما عن غيره من الموضوعات. ولم يشذ عن ذلك سوى نظام **WOW** ؛ حيث قدم مقترحات بأسماء الدول العربية كالأردن وسوريا والجزائر عند البحث باستخدام (**terrorism**).

3 / 5 البحث والاسترجاع:

تتيح نظم الاسترجاع آليات البحث ومعاملاتها المختلفة لتمكين مستخدمي النظام من وضع محددات بحث دقيقة بما يمكن للمستخدم التعبير عن احتياجاته الموضوعية بدقة كاملة تزيد من مستوى التحقيق وخفض معدل الاستدعاء في النتائج المسترجعة. ووفقا لآليات البحث تتشكل طبيعة ونوعية النتائج المسترجعة، بما يعني أن آليات البحث الأكثر تقدما وتعقيدا التي تستتر غالبا تحت مصطلح البحث المتقدم أفضل في عملية البحث من الآليات البسيطة العامة المستخدمة في النظام على واجهته الرئيسة. وتقع آليات البحث ضمن أهم اجراءات النظام المستخدمة في استرجاع نتائج ذات كيف وكم ونوع محدد، وعلى ذلك فإن النظم ذات التوجهات أو التحيز يمكنها تشكيل آليات البحث ومحدداته بما يسمح للتدخل في النتائج المسترجعة، غير أن هذه الآليات تمثل أدوات في دي المستخدم يستطيع تركيبها وهيكلتها وفقا لأغراضه المعلوماتية، ولذلك فإن إجراءات البحث داخل نظم الاسترجاع تختلف اختلافا نوعيا في قابليتها لتلبية احتياجات المستخدم بشكل مطلق أو الميل لصالح النظام في استرجاع نوعي لنتائج البحث موافقا لتوجهات وسياسات النظام. ومن ثم فإن الدراسة قد استهدفت أنواعا بذاتها من قدرات البحث وآلياته داخل نظم الاسترجاع التي إذا ما استخدمها النظام تكن لديه قابلية التحيز، أو تلك التي لا تمنح للنظام القدرة على التدخل في استرجاع موجه لنتائج البحث. ونظرا لكون هذه الاجراءات لا تحمل قيما عددية أو إحصاءات رقمية، فقد اعتمدت الدراسة على تحليل نوعية آليات وأدوات البحث المستخدمة لنظم استرجاع عينة الدراسة، واستنباط قدرتها على دعم التحيز للنظام أو توفير الحيادية لمستخدمية في النتائج المسترجعة.

جدول (7) تأثير البحث والاسترجاع بالتحيز في عينة الدراسة

إجراءات البحث والاسترجاع		نظم استرجاع الويب					
		GOOGLE	BING	YAHOO	ASK	AOL	WOW
اسلوب البحث	حر	√	√	√	√	√	√
	تصفح	---	---	---	---	---	---
مستوى البحث	الكلمات	√	√	√	√	√	√
	المفاهيم	---	---	---	---	---	---

آليات البحث	البوليني	√	√	√	√	√	√
	اللغة الطبيعية	---	---	---	---	---	---
	الدلالي	---	---	---	---	---	---
مقومات البحث	المحللات	---	---	---	---	---	---
	المعاجم	√	---	---	√	---	---
	قاعدة المعرفة	---	---	---	---	---	---

يوضح جدول (7) الجوانب التالية :-

أولا : توافقت جميع نظم العينة في نمط العمل بأداء وبرمجيات بحث بعيدة عن مستحدثات وتطورات الويب الدلالية والاسترجاع بالمفاهيم، وقد تمثل ذلك في استخدامها جميعا لأسلوب البحث الحر واستخدام الكلمات دون المفاهيم، بجانب اعتماد المنطق البوليني آلية وحيدة للبحث، بالاضافة إلى خلوها جميعا من أدوات المحللات والمعاجم اللغوية وقواعد المعرفة الداعمة لبنية الربط الدلالي بين المفاهيم.

ثانيا : لم تستخدم نظم استرجاع العينة أسلوب البحث بالتصفح رغم كونه يمثل التحكم التام والتحيز المطلق للنظام إذا ما أراد ذلك، عن طريق التحكم في ربط واسترجاع مواد الويب من خلال مصطلحات ثابتة محددة مسبقا. واستخدم النظم بدلا منها البحث الحر، وهو ما يتناسب وطبيعة مستخدمي الويب أكثر من كونه متوافق أم لا مع توجهات النظام.

ثالثا : اظهرت عمليات البحث التي اجريت عن عينة الدراسة تدني مستوى آليات البحث المتوفرة لاسترجاع مواد الويب، واستخدم فقط نظاما **Google, yahoo** قدرات البحث المعجمية في البحث عن مصطلح أجنبي **Mohamed salah** لتتضمن النتائج المسترجعة موادا باللغة العربية. ويعزي الباحث غياب القدرات المعجمية والصرفية لمفردات البحث عن نظم عينة الدراسة حتى يستطيع النظام فرض التحكم أكثر ورفع قدرة السيطرة على نوعية وطبيعة النتائج المسترجعة.

3 / 6 الفرز والترتيب للنتائج:

تحمل نتائج البحث وحدها المؤشرات الحقيقية على قياس ورصد التحيز في نظم الاسترجاع، ذلك لكونها تعتمد على الأرقام والاحصاءات التي تمثل لغة الحقائق في إجراءات وأداء نظام الاسترجاع. ويأتي ذلك من كون النتائج المسترجعة تمثل مشهد واقعي أمام مستخدمي النظام يستطيعو الحكم عليها، بل وقياسها عدديا ورقميا بما يمكن معه الوقوف على نسب تحققها بشكل مدقق. وتعتمد دراسات نظم استرجاع المعلومات على استخدام معاملات تتعلق فقط بنتائج البحث وليس الاجراءات الأخرى، فمؤشرات الاستدعاء والتحقيق **recall and precision** المستخدمة لقياس كفاءة نظام الاسترجاع تستهدف قياس النتائج بما يعكس طبيعة النظام وكفاءة أداء النظام بوجه عام. ولما كانت المحاور السابقة تعتمد على شواهد لتقييم جانب التحيز في النظام، فإن هذا المحور وغيره من المحاور التالية المتعلقة بالنتائج المسترجعة حري بها الاعتماد على صياغة معادلات ومعاملات تستخدم التوزيع العددي والإحصائي لمجموعات النتائج المسترجعة، ومن ثم الخروج بأعداد وأرقام تمثل نسبة أو

نتيجة مؤكدة لقياس درجة التحيز النوعي في أداء نظام الاسترجاع وفقا لطبيعة أعداد النتائج المسترجعة. وانطلاقا من أن النتائج المسترجعة تبدأ عند إجراء الفرز **ranking** فقد قدمت الدراسة في جدول رقم ( ) طرحاً لمعاملات لقياس التحيز في اختيار وفرز النظام لمجموعات النتائج المسترجعة ردا على استراتيجية البحث المقدمة من جانب المستخدم. وقد أوردت الدراسة هذه المعاملات في إطار الحديث عن الفرز لكون النتائج المستدعاة تكن محسومة الاسترجاع والترتيب في مرحلة الفرز قبل العرض على واجهة النتائج أمام المستخدم.

جدول (8) مقترح الدراسة لمعاملات قياس التحيزلنتائج نظم الاسترجاع

معاملات قياس التحبز	متحيز	غير متحيز
التمرير	أ	ب
الحجب	ج	

تقترح الدراسة استخدام المعاملات التالية لقياس التحيز في نتائج نظم استرجاع الويب، وهي :-

معامل التحيز (أ) = عدد النتائج المتحيزة التي تمر تمريرها ÷ إجمالي عدد النتائج = أ ÷ أ + ب

معامل الحيادية (ب)= عدد النتائج غير المتحيزة التي تمر تمريرها ÷ إجمالي عدد النتائج= أ ÷ أ + ب

معامل الحجب (ج) = عدد النتائج المحجوبة ÷ إجمالي عدد النتائج = ج ÷ أ + ب

وقد استخدمت الدراسة المعاملات السابقة للتطبيق على نتائج نظم استرجاع عينة الدراسة وفقا للجدول ( ) اعتمادا على التوزيع العددي لنتائج كلمات البحث المستخدمة مع العينة. وقد تم العمل وفقا للنتائج العشر الأولى واعتبارها جملة النتائج التي سيتم القياس عليها، مع استثناء النتائج الترويجية من القياس. استخدمت الدراسة برنامج الجداول الحسابية **Microsoft Excel** في إجراء هذا القياس ؛ حيث يتم التقاط النتائج العشر الأولى بعناوينها وتحديد قيمة الترتيب لكل نتيجة في كل نظام على حده، ثم إدارتها مع نتائج باقي نظم العينة ليتضح بالمقارنة والترتيب النتائج المميزة لكل نظام عن الباقين، وما لم يرد به وجاء بباقي النظم وهو ما يعد حجبا لهذه المواد في هذا النظام :-

جدول (9) اعداد ونسب الفرز والترتيب للنتائج في عينة الدراسة

حيث (أ) المتحيز، (ب) المحايد، (ج) المحجوب

معاملات الفرز والترتيب للنتائج		نظم استرجاع الويب											
		GOOGLE		BING		YAHOO		ASK		AOL		WOW	
		عدد	نسبة	عدد	نسبة	عدد	نسبة	عدد	نسبة	عدد	نسبة	عدد	نسبة
islam religio	(أ)	3	%30	1	%10	0	0	2	%20	3	%30	2	%20
	(ب)	7	%70	9	%90	10	100	8	%80	7	%70	8	%80

n	(ج)	3	%30	1	%10	0	0	2	%20	3	%30	2	%20
president trump	(أ)	4	%40	0	0	%20	2	5	%50	5	%50	1	%10
	(ب)	6	%60	10	100	%80	8	5	%50	5	%50	9	%90
	(ج)	4	%40	0	0	%20	2	5	%50	5	%50	1	%10
Syria	(أ)	5	%50	0	0	%10	1	1	%10	4	%40	2	%20
	(ب)	5	%50	10	100	%90	9	9	%90	6	%60	8	%80
	(ج)	5	%50	0	0	%10	1	1	%10	4	%40	2	%20
palestine israel	(أ)	1	%10	2	%20	%10	1	3	%30	3	%30	2	%20
	(ب)	9	%90	8	%80	%90	9	7	%70	7	%70	8	%80
	(ج)	1	%10	2	%20	%10	1	3	%30	3	%30	2	%20
Arab Spring	(أ)	0	0	0	0	%10	1	3	%30	2	%20	0	0
	(ب)	10	100	10	100	%90	9	7	%70	8	%80	10	100
	(ج)	0	0	0	0	%10	1	3	%30	2	%20	0	0
terrorism	(أ)	2	%20	1	%10	%10	1	2	%20	2	%20	1	%10
	(ب)	8	%80	9	%90	%90	9	8	%80	8	%80	9	%90
	(ج)	2	%20	1	%10	%10	1	2	%20	2	%20	1	%10
united nations	(أ)	2	%20	1	%10	%20	2	1	%10	1	%10	0	0
	(ب)	8	%80	9	%90	%80	8	9	%90	9	%90	10	100
	(ج)	2	%20	1	%10	%20	2	1	%10	1	%10	0	0
nelson mandela	(أ)	3	%30	1	%10	%10	1	0	0	2	%20	0	0
	(ب)	7	%70	9	%90	%90	9	10	100	8	%80	10	100
	(ج)	3	%30	1	%10	%10	1	0	0	2	%20	0	0
refuges	(أ)	2	%20	0	0	0	0	0	0	1	%10	1	%10
	(ب)	8	%80	10	100	100	10	10	100	9	%90	9	%90
	(ج)	2	%20	0	0	0	0	0	0	1	%10	1	%10

يتضح من جدول (9) الملاحظات التالية :-

أولا : اختلفت النتائج التي تحيز بها نظام ما اختلافا نوعيا عن طبيعة السياق الذي سارت عليه باقي النتائج لنفس النظام، بما يعكس تماما أن النظام تعمد إدخالها وإقحامها بين نتائج البحث الأخرى التي توافقت بقدر كبير مع باقي نتائج النظم الاخرى. وذلك لما لاحظه الباحث من طبيعة المحتوى الذي تتعرض له هذه النتائج بشكل جلي في طبيعة الرسالة التي تريد ايصالها.

ثانيا : إن النتائج التي تميز بها كل نظام على حده لم تكن بواقع اختلاف حجم ونوعية مواد المعلومات التي يملكها، بل عن إرادة تحيز لصالح هذه المواد وذلك لكون النظم الآن تقترب في ظل تطور تقنياتها

من قدرات عالية في الاضافة والمعالجة وتماثل الاجراءات والأدء يمكنها من تقارب شديد جدا في نوعية وحجم مواد المعلومات.

ثالثا : تأثرت النتائج بخاصية البحث في الوقت الحقيقي **real time search** كما في **Google** حيث تميزت به عن غيرها، كما انعكس تأثير العنصر البشري في بعض النظم مثل **WOW, AOL** حيث كانت النتائج الأولى فيها الأكثر دقة والأقرب لطبيعة البحث فمثالا ازهر الموضع الرسمي للرئيس الامريكي ترمب عند البحث باسمه، بخلاف باقي نظم العينة.

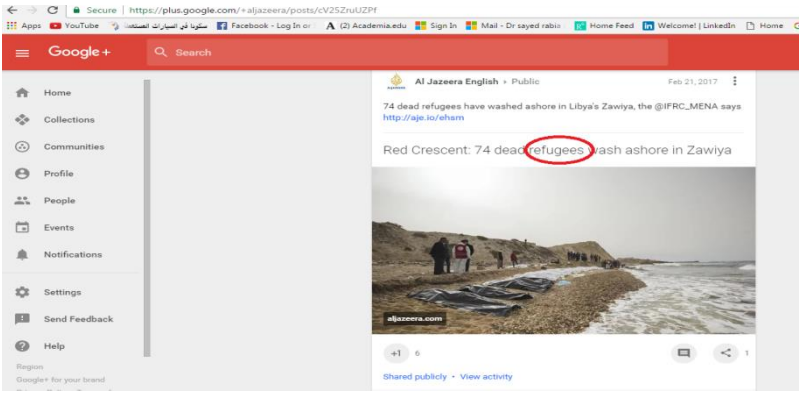
رابعا : منحت القدرات اللغوية لبعض النظم القدرة على التميز ببعض النتائج المحتلفة لغويا عن مصطلحات البحث كما في نظام **google** ؛ حيث تمكن النظام من استرجاع نتائج نوعية لا توجد في باقي عينة الدراسة. فمن ناحية يمكن اعتبار ذلك ملمحا ايجابيا لنظام الاسترجاع في إعلاء الدقة على مستوى الاسترجاع بشكل لغوي محدد، ومن ناحية أخرى فهو قدر أكبر من التحكم في عرض النتائج التي يفضلها النظام عن غيرها. وهو ما سبب ارتفاع عدد النتائج التي تميز بها **google** عن غيره مع مصطلحي **Syria, Nelson Mandela**.

خامسا : لقد أوضحت اعداد النتائج المسترجعة مدى التقارب في أعداد نتائج التحيز في كل من نظم العينة **WOW, ASK, AOL** من جانب واختلافهما عن نظامي **yahoo, google** من جانب آخر. ويعزي الباحث ذلك لطبيعة الأداء في هذه النظم الأقرب للاشراف البشري على أدارة مجموعات النظام، بما يختلف بدوره عن الاداء الآلي كلية في نظامي **google, yahoo**.

سادسا : لم تعكس نتائج البحث تباينا كبيرا بين نظم عينة الدراسة عند استخدام مصطلح **arab spring**، ما وصل لدرجة توحد ثلاثة نظم **google, bing, wow** في النتائج المسترجعة وعدم تميز أي منها عن النظم الأخرى. وذلك يعزيه الباحث إلى اختلاف هذه القضية عن سابقاتها من مصطلحات البحث الأخرى في كونها قضة تتعلق بطرف واحد هو العالم العربي، بعكس القضايا التي يدخل فيها العالم الغربي موطن نظم الاسترجاع كطرف فيها، ومن ثم تكن له توجهات أو تداخلات تنعكس بدورها على نتائج البحث.

سابعا : كادت نتائج البحث ان تتوافق بين نظم استرجاع عينة الدراسة حين البحث عن مصطلح يمثل موضوعا محايدا بطبعه يخلو من أي استقطابات أو توجهات هو **united nations** ؛ حيث لا تمثل هذه المؤسسة قضية متعددة الأطراف أو التوجهات بما يجعلها موضعا لتنازع الرؤى والآراء بما ينعكس على تنوع وتباين النتائج المسترجعة لكل نظام على حده. حيث ظهر ذلك التوافق جليا بين نظم **ASK, AOL, BING** من جهة وبين نظامي **yahoo, google** من جهة أخرى.

ثامنا : دللت نوعية النتائج التي تميزت بها نظم الاسترجاع صحة الفرضية الخاصة بتحيز النظام لمواد المعلومات التي تنشر على موقعه وترجيح قيمتها أكثر من غيرها من النتائج الأخرى، وقد بدا ذلك جليا في نظام الاسترجاع **google** حين البحث بمصطلح **refugees** بإدراجه لمنشورات **posts** حول هذا الموضوع نشرت على واجهة **plus.google** إحدى خدمات **google**. وذلك كما يوضح الشكل (6) بصحة تحيز نظم الاسترجاع وتدخلها في طبيعة ونوعية النتائج المسترجعة.



شكل (6) تحيز نظام google في استرجاع مواد المعلومات المنتمية لمجال نشره

3 / 7 عرض النتائج:

ترتبط النتائج المعروضة بإجراءات الفرز المستخدمة في نظام الاسترجاع وهي ما ينعكس عليها نسبة ودرجات التحيز المنتهجة من جانب النظام بتمرير أو حجب النتائج، وهو ما يمكن أن نطلق عليه تحيزا نوعيا من جانب النظام في اختيار مواد المعلومات التي سيشاهدها المستخدم. غير أن هذا التحيز النوعي يرتبط به تحيز آخر هو التحيز الشكلي في اختيار قيمة ومكانة كل نتيجة للرد على استراتيجية البحث ؛ حيث تعمد نظم استرجاع الويب إلى ترتيب النتائج وفقا لآليات محددة تأخذ من خلالها مادة المعلومات قيمة عالية أو منخفضة بين النتائج معبرة عن مدى صلة مادة المعلومات بموضوع البحث وأهميتها في تلبية الحاجات المعلوماتية للمستخدم. وقد طرحت الدراسة معاملين متكاملين هما الحياد والتحيز لقياس مدى توافق النظام مع باقي نظم الاسترجاع عند استخدام استراتيجية بحث ومفردات محددة للبحث، وهو ما يعكس على أثره تحيز النظام مع مواد معلومات محددة بخفض أو رفع قيمتها للتعديل والتغيير في شكل النتائج المسترجعة، ومن ثم التحيز وفرض توجه بتقديم نتائج على أخرى أو تأخير نتائج قد لا يرها المستخدم، خاصة في ظل أعداد مليونية من نتائج البحث على واجهات نظم استرجاع الويب. ويتلخص تطبيق التحيز في أجرائين ؛ أولهما : تحديد متوسط القيمة المكانية لترتيب كل مادة معلومات بذاتها في مجموعة من نظم الاسترجاع المكافئة أو المناظرة باستخدام معامل الحياد ، ثانيهما : قياس قيمتها داخل نظام الاسترجاع موضع التقييم، وبالتالي ينتج مدى تحيز أو حيادية النظام مع هذه المواد باستخدام معامل التحيز .

أولا : معامل الحياد = متوسط قيمة الوحدة بين النظم ÷ عدد نتائج المعروضة للنظم المقارنة
اعتمد القياس على النتائج التي وردت في كل نظم العينة حتى تتكون صورة كلية لتباين الترتيب بينها جميعا، وتحديد فوارق الاهتمام بذات النتيجة في مختلف نظم العنية. ومُنحت النتائج نقاطا بحسب قيمتها في ترتيب الاسترجاع ؛ حيث تأخذ النتيجة 10 نقاط إذا كانت أولى النتائج العشر المسترجعة موضع القياس، وتأخذ النتيجة الثانية 9 نقاط، وهكذا حتى تاخذ النتيجة العاشرة نقطة واحدة.

جدول (10) اعداد ونسب معامل الحياد للنتائج في عينة الدراسة

معامل الحياد		نظم استرجاع الويب						
مصـــــــطلح البحث	عنوان النتيجة المسترجعة	GOOGLE	BING	YAHOO	ASK	AOL	WOW	معامل الترتيب المحايد
islam religion	Islam - ReligionFacts	**8**	8	8	7	8	6	**8**
president trump	President Trump (@POTUS) | Twitte	6	7	8	7	6	2	**6**
Syria	Syria - Wikipedia	10	10	10	8	10	10	**10**
palestine Israel	Primer on Palestine, Israel and the Arab-Israeli Conflict	6	7	7	7	8	6	**7**
Arab Spring	Home - Arab Spring: A Research & Study Guide	5	7	5	5	7	5	**6**
terrorism	Terrorism — FBI	8	9	9	9	7	8	**8**
nelson mandela	Biography – Nelson Mandela Foundation	7	8	8	9	5	8	**8**
Refugees	UNHCR - Refugees	8	7	8	6	8	7	**8**

ثانيا : معامل التحيز = (قيمتها في النظام - معامل الترتيب المحايد)

جدول (11) اعداد ونسب معامل التحيزللنتائج في عينة الدراسة

معامل الحياد			نظم استرجاع الويب					
مصـــــــطلح البحث	عنوان النتيجة المسترجعة		GOOGLE	BING	YAHOO	ASK	AOL	WOW
islam religion	Islam - ReligionFacts	قيمتها بالنظام	**8**	8	8	7	8	6
		معامل الحياد	**8**	**8**	**8**	**8**	**8**	**8**
		معامل التحيز	**0**	**0**	**0**	**1-**	**0**	**2-**
		مؤشر التحيز	محايد	محايد	محايد	سلبي	محايد	سلبي
president trump	President Trump (@POTUS) | Twitte	قيمتها بالنظام	6	7	8	7	6	2
		معامل الحياد	**6**	**6**	**6**	**6**	**6**	**6**
		قيم التحيز	**0**	**1+**	**2+**	**1+**	**0**	**4 -**
		مؤشر التحيز	محايد	متحيز	متحيز	متحيز	محايد	سلبي

Syria	Syria - Wikipedia	قيمتها بالنظام	10	10	10	8	10	10
		معامل الحياد	**10**	**10**	**10**	**10**	**10**	**10**
		قيم التحيز	**0**	**0**	**0**	**2-**	**0**	**0**
		مؤشر التحيز	محايد	محايد	محايد	سلبي	محايد	محايد
palestine Israel	Primer on Palestine, Israel and the Arab-Israeli Conflict	قيمتها بالنظام	6	7	7	7	8	6
		معامل الحياد	**7**	**7**	**7**	**7**	**7**	**7**
		قيم التحيز	**1-**	**0**	**0**	**0**	**1+**	**1-**
		مؤشر التحيز	سلبي	محايد	محايد	محايد	متحيز	سلبي
Arab Spring	Home - Arab Spring: A Research & Study Guide	قيمتها بالنظام	5	7	5	5	7	5
		معامل الحياد	**6**	**6**	**6**	**6**	**6**	**6**
		قيم التحيز	**1-**	**1+**	**1-**	**1-**	**1+**	**1-**
		مؤشر التحيز	سلبي	متحيز	سلبي	سلبي	متحيز	سلبي
Terrorism	Terrorism — FBI	قيمتها بالنظام	8	9	9	9	7	8
		معامل الحياد	**8**	**8**	**8**	**8**	**8**	**8**
		قيم التحيز	**0**	**1+**	**1+**	**1+**	**1-**	**0**
		مؤشر التحيز	محايد	متحيز	متحيز	متحيز	سلبي	محايد
nelson mandela	Biography – Nelson Mandela Foundation	قيمتها بالنظام	7	8	8	9	5	8
		معامل الحياد	**8**	**8**	**8**	**8**	**8**	**8**
		قيم التحيز	**1-**	**0**	**0**	**1+**	**3-**	**0**
		مؤشر التحيز	سلبي	محايد	محايد	متحيز	سلبي	محايد
Refugees	UNHCR - Refugees	قيمتها بالنظام	8	7	9	8	6	7
		معامل الحياد	**8**	**8**	**8**	**8**	**8**	**8**
		قيم التحيز	**0**	**1-**	**1+**	**0**	**2-**	**1-**
		مؤشر التحيز	محايد	سلبي	متحيز	محايد	سلبي	سلبي

يتضح من الجدول (11) المشاهدات التالية :-

أولا : فرضت طبيعة محتوى النتيجة المسترجعة تأثيرها على موضع ترتيب وتقييم النتيجة بين جملة النتائج المسترجعة ؛ حيث تباينت قيم النتائج التي تتسم بوجهات النظر المتنوعة بين المحايد والسلبي والمتحيز، نظرا لما لها من اختلاف في المعالجة والتناول. في حين أن النتائج أحادية التناول مثل الحديث عن سوريا في الموسوعة الحرة فإن تناولها لن سيكون أحاديا لأنها حقائق عن دولة بشكل مجرد وبالتالي لا يؤثر تقديمها أو تأخيرها بين النتائج المسترجعة.

ثانيا : أوضحت قيم الترتيب للنتائج أنه لا يوجد نظام واحد بين نظم العينة قد ثبت في أدائه على نحو محدد بالتحيز أو الحياد ؛ وذلك يعزى لاختلاف طبيعة النتائج بين عملية بحث وأخرى تبعا لاختلاف

القضايا موضع البحث. بالاضافة إلى أن هناك عوامل اخرى مثل قدرات المعالجة اللغوية والفنية لمواد المعلومات المختلفة في مستواها بين نظام وآخر.

جدول (12) تأثير عرض النتائج بالتحيز في عينة الدراسة

إجراءات عرض النتائج	نظم استرجاع الويب					
	GOOGLE	BING	YAHOO	ASK	AOL	WOW
توزيع النتائج	√	√	√	√	√	√
تقسيم النتائج	√	√	√	√	√	√
تقديم نتائج مقترحة	√	---	√	√	√	√
عرض كلمات البحث مع النتائج	√	√	√	---	---	√

يتضح من الجدول (12) أن توزيع النتائج على واجهات نظم الاسترجاع يعكس توجهات النظام في التعامل مع مواد المعلومات والمجالات الموضوعية مسار البحث ؛ حيث وزعت النظم موادها الإعلانية في بداية عرض النتائج، حتى تكون أول ما تقع عليه أعين مستخدميها بما يزيد مقابل الإعلانات. أيضا فإن تقسيم واجهة عرض النتائج يفيد في إعلاء درجة الحيادية بما يجعل المستخدم قادرا على تصفح نتائج البحث التي تخص الموضوع بعيدا عن المؤثرات الاخرى. وتنتهج بعض النظم منهج إبراز كلمات البحث بنمط كتابة مختلف في عرضها ضمن السياق الذي وردت فيه، وهو ما يدعم خصائص الحياد للنظام في تأكيد مصداقية اختيار واسترجاع هذه النتائج.

المبحث الرابع رؤى ونتائج الدراسة:

1) **أن** نظم استرجاع الويب لم تعد نظم المعالجة والبحث لمواد الويب بشكل محايد من منظور دراسات الإفادة في مجال المعلومات، بل ينظر الآن لنظم الاسترجاع باعتبارها نظم إدارة وتداول المعرفة الرقمية بما يوافق سياسات عمل ممنهجة. فلا يقتصر الأمر الآن في نظام الاسترجاع على اتاحة مادة الويب وفقا لما هي عليه، إنما يتعدى ذلك إلى تخير موضع وموضوع الإتاحة وصب المحتوى في معين محدد. ولذلك **يجب** أن يتغير التناول العلمي والبحثي لنظم استرجاع الويب في دراستها، بما يوافق هذه الحقيقة حتى يساهم في تحقيق الإفادة الفعلية لمستخدمي الويب.
2) **أن** هناك تعاطي غير خفي الآن في بيئة الويب بين عمليات النشر وبث المعلومات وإتاحتها، وتمثل ذلك في مواد الويب النوعية المحررة بأهداف أيديولوجية وسياسية وإقتصادية وتداخلها مع مواد الويب في عمليات المعالجة والاسترجاع، ومن ثم حملت مواد الويب في نشرها جين التحيز الذي انتقل بدوره إلى نظم الاسترجاع مانحا دعما أكبر للتدخل والتحكم من النظام وإعلاء توجهاته وسياساته الاقتصادية والسياسية في بحث واسترجاع مواد الويب. لذلك **يجب** نشر وتنمية الوعي البحثي وخصائص التقييم والنقد لدى مستخدمي مواد معلومات الويب ونظم استرجاعها.
3) **أن** قدرات المعالجة اللغوية والدلالية لنظم الاسترجاع منحتها درجة متقدمة في مستوى الربط الثقافي والمزج المعرفي بين بيئات النشر للمحتوى العربي وغيره من الثقافات الأخرى، مما جعل نظم

استرجاع الويب قادرة على حمل لقاحات الثقافات الأخرى للمستخدم العربي بل والتدخل في تحديد توجهاته البحثية ودعم نشر أو تقييد وجهة فكرية ما. ولذلك **يجب** أن يتسم المحتوى العربي بالقدرة على التعبير عن توجهاته وقضاياه المختلفة التي تتخذ نظم استرجاع الويب الآن منصة للتعبير عنها وإطلاقها إلى مستخدمي الويب.

4) **أن** نظم الاسترجاع لا تعمل في بيئة منعزلة عن مجريات وقضايا الواقع الانساني رغم كونها تعمل في بيئة الواقع الافتراضي، وهو ما يعزي وجود توجهات ومواضع تحيز من النظم في استرجاع مواد الويب في جوانب سياسية أو أيديولوجية أو فكرية بذاتها. وقد مزجت شركات نظم الاسترجاع بين الطابعي الاقتصادي والطابع السياسي والطابع المعلوماتي في قضية بحث واسترجاع مواد الويب. ولذلك **يجب** إضافة أبعاد أخرى لتقييم نظم استرجاع الويب وأدائها بالاعتماد على تحليل توجهات الكيانات المالكة أو الراعية لهذه النظم.

5) **أن** طبيعة الويب وانتشارها على مستوى المستخدمين بفئاتهم واهتماماتهم النوعية فرضت تحويل الويب إلى بيئة تنافسية في مجالات السياسة والانتخابات الحزبية والرئاسية، فضلا عن المنافسات التجارية وتوحويلها إلى بيئة خصبة لنشر التوجهات والأفكار المختلفة. وقد انعكس ذلك بآثاره على خدمات البحث والاسترجاع لمواد الويب المتلونة مسبقا بألوان التحيز في نشرها، بل السعي الحثيث لتغليب وجهات نظر على أخرى بتغليب إتاحة مواد معلومات على أخرى في عمليتي النشر والإتاحة. ولذلك **يجب** مراعاة توقيتات وأماكن النشر لمواد معلومات الويب ونطاق نشرها، خاصة إذا ما تلونت بألوان الأحداث الانتخابية والسياسة في منطاق وقوعها.

6) أثبتت الدراسة التطبيقية **أن** نتائج البحث وحدها تمثل المقياس الصادق في تحديد ملامح التحيز مستوياته وأشكاله في نظم استرجاع الويب، وذلك لكونها المعبرة عن المنتج النهائي المستهدف إفادة وتوصيل مفاهيم ومعلومات محددة لمستخدم النظام. ولذلك **يجب** الاهتمام بتطوير ودعم القياسات الإحصائية والعديدة في التعامل مع نظم استرجاع الويب، وقد قدمت الدراسة طرحا لمجموعة معاملات ومؤشرات لقياس التحيز والحيادية في نظم الاسترجاع.

7) أثبتت الدراسة التطبيقية صحة الفرضية الأساسية التي بنيت عليها هذه الدراسة **بأن** نظم استرجاع الويب تصطبغ بملامح تمثل توجهاتها النوعية في التعامل مع مواد الويب ومستخدميها بما ينعكس ذلك على نوعية وطبيعة وتريب النتائج المسترجعة، بما يعني التدخل المتعمد من النظام في إدارة النتائج وفقا لرغبة محددة في آلية وطرائق العرض أكثر من كونها آليات بحث واسترجاع تقنية وبرمجيات تعمل على المعالجة المحايدة.

8) **أن** نظم استرجاع الويب تتأثر حتما بطبيعة التوجهات التي بنيت على أثرها او الثقافة التي تخدمها وينعكس ذلك حتما على طبيعة النتائج والمعلومات المقدمة لمستخدميها، ومن ثم فإن عامل التحيز لمحركات البحث العربية قد لا يمكن اغفاله في نظم استرجاع المعلومات العربية، ذلك لما تتسم به البيئة العربية من تجاذبات سياسية واجتماعية وفئوية خاصة. ولذلك **يجب** الشروع في وضع

تقانين محددة لسياسات وآليات بناء وعمل نظم استرجاع المعلومات العربية التي تتعامل مع محتوى عربي ثري بمواد إعلامية وعائية ومعلوماتية نوعية شتى.

9) أن التطورات التقنية للنشر الرقمي في بيئة الويب استطاعت إخراج مستخدم الويب من سيطرة الرأي الواحد والتوجهات المقننة في الرأي والمعرفة، وسبب انفتاح بيئة النشر الرقمي التعبير عن الذات. وقد عملت التقنيات والتطورات البرمجية مرة أخرى على تضييق قدرة نظم استرجاع الويب على فرض توجهاتها وقللت من قدرة النظم على التحكم المطلق في نتائج البحث المسترجعة من خلال تقنيات الويب الدلالي التي أخرجت المفاهيم والمعاني من ضيق اللفظ والكلمة إلى سعة قواعد المعرفة، بما يصعب على نظام الاسترجاع التحكم في المعالجة والنتائج بدرجة كبيرة.

1. **Moshowitz, abbe and kawaguchi, akira. Assessing bias in search engines, information processing and management, 2002, cited at 25/10/2016, cited at http://p9797-ugrade2.eul.edu.eg.ugrade1.eul.edu.eg:2048/MuseSessionID=0004ity/MuseProtocol=http/MuseHost=www.sciencedirect.com/MusePath/science/article/pii/S0306457301000206/pdfft?md5=21a115d66855e0ae0d0810ad3ce6b852&pid=1-s2.0-S0306457301000206-main.pdf**
2. **Wright, Joshua D. Defining and Measuring Search Bias: Some Preliminary Evidence, George Mason University School of Law, 2011, cited at 25/10/2016, cited at https://www.law.gmu.edu/assets/files/publications/working_papers/1214DefiningandMeasuringSearchBias.pdf**
3. **Tedla, B.E. Suman. ANALYZING BIAS AND QUALITY OF SEARCH ENGINES USING HITS, UMI ProQuest Information and Learning Company, 2007, cited at 25/10/2016, cited at http://p9797-ugrade2.eul.edu.eg.ugrade1.eul.edu.eg:2048**
4. **Moshowitz, abbe and kawaguchi, akira. OP. CT.**
5. **Moshowitz, abbe and kawaguchi, akira. Measuring search engine bias, Information Processing and Management 41 (2005) 1193–1205, cited at 25/10/2016, cited at www.sciencedirect.com/MusePath/science/article/pii/S0306457304000615/pdfft?md5=2130ae43b6cc9de3e28968c3476f8549&pid=1-s2.0-S0306457304000615-main.pdf**
6. **BIAS ON THE WEB. COMMUNICATIONS OF THE ACM September 2002, cited at 1/9/2016, cited at http://cacm.acm.org/magazines/2002/9/20780-bias-on-the-web/fulltext**
7. **Chen, Ing-Xiang and Yang, twCheng-Zen. Position Paper: A Study of Web Search Engine Bias and its Assessment, IW3C2 WWW 2006, cited at 20/10/2016, cited at http://www.olmedilla.info/events/MTW06_papers/paper08.pdf**

8. E. Goldman. Search Engine Bias and the Demise of Search Engine Utopianism, Springer-Verlag Berlin Heidelberg 2008, cited at 11/9/2016, cited at link.springer.com/MusePath/content/pdf/10.1007%2F978-3-540-75829-7.pdf
9. Internet Live Stats. World wide web consortium, cited at 1/11/2016: 2:27 pm, cited at http://www.internetlivestats.com/
10. Owens, Ciaran. A Study of the Relative Bias of Web Search Engines Toward News Media Providers, University of Glasgow, 2009, cited at 13/12/2016, cited at https://pdfs.semanticscholar.org/bd99/dc4159a9f821e5fcf9f58aaae166899da495.pdf
11. Kąkol, Michał and Jankowski-Lorek, Michał. On the Subjectivity and Bias of Web Content Credibility Evaluations, international world wide web conference committee, 2005, cited at 11/11/2016, cited at https://infoscience.epfl.ch/record/196260/files/p1131.pdf
12. The Commercial Search Engine Industry and Alternatives to the Oligopoly. Interactive Media Studies and Journalism, Miami University of Ohio, 2005, cited at 20/10/2016, cited at http://www.mediaweek.com/mw/news/interactive/article_display.jsp?vnu_content_id=1000096570
13. Diaz, Alejandro M..Advertising and "Mixed Motives" : Exploring Search Engine (Hyper)commercialization, 2005, cited at 12/11/2016, cited at https://pdfs.semanticscholar.org/efaa/88199621af0bea3bc4cf26127731e81a82fd.pdf
14. The Commercial Search Engine Industry and Alternatives to the Oligopoly. OP. CT.
15. Diaz, Alejandro M.. Through the Google Goggles:Sociopolitical Bias in Search Engine Design, the author, Stanford University, 2005, PHD, cited at 11/11/2016, cited at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.457.8268&rep=rep1&type=pdf#page=18
16. Ibid.
17. Jansen, Bernard J. and Spink, Amanda H. Sponsored Search: Is Money a Motivator for Providing Relevant Results?. IEEE, 2007, cited at 11/9/2016, cited at http://eprints.qut.edu.au/archive/00014139/01/14139.pdf.
18. Couvering, Elizabeth Van . New Media? The Political Economy of Internet Search Engines, Conference of the International Association of Media & Communications Researchers (IAMCR) Porto Alegre, Brazil, July 25-30, 2004, cited at 10/10/ 2016, cited at

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.129.1900&rep=rep1&type=pdf

19. Feng, Juan and Bhargava, Hemant K. Implementing Sponsored Search in Web Search Engines: Computational Evaluation of Alternative Mechanisms, Knowledge and Data Management, 2005, cited at 15/11/2016, cited at https://www.researchgate.net/profile/Juan_Feng6/publication/220669588_Implementing_Sponsored_Search_in_Web_Search_Engines_Computational_Evaluation_of_Alternative_Mechanisms/links/574811a408ae2301b0b8db4d.pdf
20. Kaid, Lynda Lee. Political Advertising and InfoiTaation Seeking: Comparing Exposure via Traditional and Internet Channels, Journal of Advertising, Taylor & Francis, Ltd. 2002, cited at 14/12/2016, cited at http://www.jstor.org/stable/4189205
21. Pew Research Center, November 13, 2008; Pew Internet and American Life Project, March 17, 2011, cited at 15/6/ 2016, cited at http://www.pewinternet.org/2011/03/17/the-internet-and-campaign-2010/
22. Ibid
23. Epstein, Robert and Ronald E. Robertson. The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections, American Institute for Behavioral Research and Technology, 2014, cited at 14/10/2016, cited at www.pnas.org/cgi/doi/10.1073/pnas.1419828112
24. Ibid.
25. Mustafaraj, Eni and Metaxas, Panagiotis (2010). "From Obscurity to Prominence in Minutes: Political Speech and Real-Time Search." In: Proceedings of the WebSci10: Extending the Frontiers of Society On-Line, April 26-27th, 2010, Raleigh, NC: US.
26. Ulken, Eric. A Question of Balance: Are Google News search results politically biased? A QUESTION OF BALANCE, 2005, cited at 13/9/2016, cited at http://ulken.com/thesis/googlenews-bias-study.pdf
27. Owen, Diana. The Internet and Voter Decision-Making, Georgetown University, 2011, cited at 13/12/2016, cited at http://www.democracy.uci.edu/files/docs/conferences/2011/Internet.VoterDecisionMaking-Diana%20Owen.pdf
28. Settle, Jaime E. and Bond, Robert and Coviello, Lorenzo. From Posting to Voting: The Effects of Political Competition on Online Political Engagement, University of California, San Diego, Department of Political Science, 2014, cited at cited at

13/9/2016, cited at https://pdfs.semanticscholar.org/ff66/420616267a69fddcde9dae6aceacd676c894.pdf

29. Scherer, Aaron. The impact of political cues on information seeking and the need for cognitive closure. PhD (Doctor of Philosophy), thesis, University of Iowa, 2014, cited at 13/9/2016, cited at http://ir.uiowa.edu/etd/4747.
30. Udoka, Udensi Eugene. Social Media and Political Effects: A Case Study of the 2015 Nigeria's Presidential Election, International Journal of Social Science and Humanities Research, 2015, cited at 14/10/2016, cited at http://www.researchpublish.com/download.php?file=Social Media and Political Effects-1488.pdf&act=book
31. Lev-On, Azi. The Democratizing Effects of Search Engine Use: On Chance Exposures and Organizational Hubs, University of Pennsylvania, 2009, cited at 13/9/2016,cited at http://ssrn.com/abstract=1481901
32. Marett, Alexandra. Participating Online: The Internet and its Role in Political Participatory Behaviour in the Context of the New Zealand General Election 2008, University of Canterbury, 2010, cited at 14/10/2016, cited at http://ir.canterbury.ac.nz/bitstream/handle/10092/4962/thesis_fulltext.pdf?sequence=1
33. Benjamin Edelman. BIAS IN SEARCH RESULTS?: DIAGNOSIS AND RESPONSE, THE INDIAN JOURNAL OF LAW AND TECHNOLOGY VOLUME 7, 2011, cited at 11/11/ 2016, cited at http://www.ijlt.in/archive/volume7/2_Edelman.pdf
34. Lewandowski, Dirk. Credibility in Web Search Engines, Hamburg University of Applied Sciences, 2012, cited at 2/5/2016, cited at https://arxiv.org/pdf/1208.1011
35. Farag, Mohamed M. G. and Lee, Sunshin and Fox, Edward A. Focused crawler for events, Springer-Verlag Berlin Heidelberg, 2017, cited at 1/1/2017, cited at link.springer.com/MusePath/article/10.1007/s00799-016-0207-1
36. Ibid.
37. Sun, Yang. A COMPREHENSIVE STUDY OF THE REGULATION AND BEHAVIOR OF WEB CRAWLERS, the author, 2008, cited at 13/12/2016, cited at https://etda.libraries.psu.edu/files/final_submissions/6784
38. Tomasz Ku´smierczyk1 and Marcin Sydow, Towards a Keyword-Focused Web Crawler, Springer-Verlag Berlin Heidelberg , 2013, cited at 13/12/2016, cited at http://students.mimuw.edu.pl/~tk290810/works/crawler.pdf

39. Andersson, Fredrik K. and Silvestrov, Sergei D. The Mathematics of Internet Search Engines, Springer Science+Business Media B.V. 2008, cited at 13/12/2016, cited at http://link.springer.com/article/10.1007/s10440-008-9254-y
40. Demner-Fushman, Dina and Shooshan, Sonya E. Annotation of Chest Radiology Reports for Indexing and Retrieval , Springer International Publishing Switzerland 2015, , cited at 18/1/2017, cited at http://link.springer.com/chapter/10.1007/978-3-319-24471-6_9
41. ElAraby, M. E. and Sakre, M. M. Dynamic and Distributed Indexing Architecture in Search Engine using Grid Computing, International Journal of Computer Applications, 2012, cited at 10/1/2017, cited at https://pdfs.semanticscholar.org/8155/4aa22821acaad21957449a86b4d8e10645d1.pdf
42. Minnie, D. and Srinivasan, S. Multi-Domain Meta Search Engine with an Intelligent, Springer-Verlag Berlin Heidelberg, 2011, cited at 1/2/2017, cited at http://link.springer.com/chapter/10.1007/978-3-642-24043-0_13

Interface for Efficient Information Retrieval on the Web

43. Kowalski, G. Information Retrieval Architecture and Algorithms, Springer Science+Business Media, LLC, 2011, , cited at 1/1/2017, cited at download.springer.com/MusePath/static/pdf/758/bok
44. Lewandowski, Dirk. Is Google Responsible for Providing Fair and Unbiased Results? Springer, 2017, , cited at 12/2/2017, cited at http://link.springer.com/chapter/10.1007/978-3-319-47852-4_4
45. Bozdag, Engin. Bias in algorithmic filtering and personalization, Springer, 2013, cited at 10/1/2017, cited at http://link.springer.com/article/10.1007/s10676-013-9321-6
46. Sullivan, Danny. How Google Instant's Autocomplete Suggestions Work, searchengineland, 2011, cited at 1/1/2017, cited at http://searchengineland.com/how-google-instant-autocomplete-suggestions-work-62592
47. Ibid.
48. Information Retrieval and Web Search, Springer, 2011, cited at 1/1/2017, cited at http://www.langtoninfo.co.uk/web_content/9780521865715_frontmatter.pdf
49. Clear. Brendan & Exton. Chris & Buckley, Jim. An empirical analysis of information retrieval based concept location techniques in software comprehension, Springer, 2008, , cited at 12/2/2017, cited at http://www.academia.edu/download/46950702/s10664-008-9095-320160702-24471-8nv7k0.pdf

50. Piwowar, Heather A and Chapman,Wendy W. Recall and bias of retrieving gene expression microarray datasets through PubMed identifiers, Journal of Biomedical Discovery and Collaboration, 2010, cited at 14/10/2016, cited at http://www.ojphi.org/ojs/index.php/jbdc/article/viewFile/2785/2481
51. Nekrestyanov, I. S. and Panteleeva, N. V. Text Retrieval Systems for the Web, Programming and Computer Software,2002, cited at 12/2/2017, cited at link.springer.com/MusePath/article/10.1023/A%3A1016371117213
52. Hawking, David. Very Large Scale Information Retrieval, Springer, 2003, cited at 12/3/2017, cited at link.springer.com/MusePath/chapter/10.1007/978-3-540-45115-0_5
53. Koopman, Bevan and Zuccon , Guido And Bruza, Peter. Information retrieval as semantic inference: a Graph Inference model applied to medical search, Springer, 2015, cited at 13/12/2016, cited at http://link.springer.com/article/10.1007/s10791-015-9268-9
54. Stoyanovich, Julia and Mee, William and Ross, Kenneth A. Semantic Ranking and Result Visualization for Life Sciences Publications, Columbia University, 2011, cited at 12/2/2017, cited at http://www.cs.columbia.edu/~kar/pubsk/ICDE10_conf_full_040.pdf
55. Bouramoul, Abdelkrim and Kholladi, Mohamed-Khireddine. An ontology-based approach for semantic ranking of the web search engines results, University of Mentouri Constantine, 2012, , cited at 1/1/2017, cited at https://arxiv.org/pdf/1212.2587
56. Zhao, Fang and Huang, Yongzhen. Deep Semantic Ranking Based Hashing for Multi-Label Image Retrieval, IEEE explorer, 2015, cited at 12/1/2017, cited at http://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Zhao_Deep_Semantic_Ranking_2015_CVPR_paper.html
57. Bashir, Shariq Rauber, Andreas. Automatic ranking of retrieval models using retrievability measure, Springer, 2014, cited at 16/2/2017, cited at link.springer.com/MusePath/article/10.1007/s10115-014-0759-6
58. Akritidi, Leonidas. Engineering Search Algorithms for Web Data, University of Thessaly, 2013, cited at 12/2/2017, cited at http://www.e-ce.uth.gr/wp-content/uploads/formidable/Akritidis_leonidas.pdf
59. Alexa. Web ranking 2017, Cited at 12/2/2017, cited at http://www.alexa.com/topsites/category/Computers/Internet/Searching/Search_Engines