

البحث باللغة العربية على محرك البحث جوجل (Google)

نبيل بن عبدالرحمن المعثم*

١ / ٠ تمهيد:

عملت الإنترنت على فتح آفاق التواصل العلمي من خلال حركة إنتاج المعلومات الرقمية ونشرها على الويب في مختلف المجالات الموضوعية ومختلف الوسائط المتعددة التي تحملها هذه البيئة الرقمية. وبالإضافة إلى ذلك فقد استطاعت مختلف الثقافات واللغات المتداولة في العالم رسم خريطة تواجه فيها هذا الكم الهائل

غير أن الويب ومحركات البحث على الإنترنت افتقرت في مدد طويلة إلى التعامل مع مواد المعلومات المكتوبة باللغة العربية، وشاب هذا التعامل النقص والافتقار إلى برمجيات وآليات البحث الداعمة للتعامل مع اللغة العربية. ويحاول الباحث في هذه الدراسة إلقاء الضوء على خصائص اللغة العربية وطبيعة المتن العربي على الويب، وطبيعة تعامل

من المعلومات الرقمية، ومن بين هذه اللغات كانت اللغة العربية بكل ما تحمله من أفكار ومعلومات متمثلة في مصادر المعلومات العلمية والمواقع التابعة لهيئات ذات أنشطة واتجاهات مختلفة. ومن ثم فقد طرحت اللغة العربية كمّاً هائلاً من المعلومات الرقمية أمام المستفيدين الذين يصلون إلى هذا الكم ويتداولونه بلغة كتابته المكونة من الحروف العربية.

- * بكالوريوس علم المكتبات والمعلومات من جامعة الملك سعود، ١٩٩٥م.
- ماجستير علوم المكتبات من معهد برات - نيويورك بأمريكا، عام ٢٠٠٢م.
- دكتورة في المكتبات والمعلومات من جامعة أم درمان الإسلامية ٢٠١٠م.
- يعمل حالياً مديراً لإدارة البحوث والنشر بمكتبة الملك فهد الوطنية.

يُسمّى باقتصاد المعرفة (Knowledge Economy). وتنتج جميع الدول العربية إلى تبني تقنيات التعليم الإلكتروني في تطوير أنظمتها التعليمية^(١).

وتنتج دور إنتاج مصادر المعلومات الرقمية العربية إلى سد الفجوة الرقمية، واللحاق بالركب العالمي حيث تتزايد مواد المعلومات العربية الرقمية على الويب باطراد مستمر، ومع التطور العلمي والتقني، وتطوير مهنة النشر الإلكتروني لمواد المعلومات المحسبة بصفة خاصة، بدأ الناشر العربي بالتحول إلى تنوع طرق النشر لتشمل النشر الورقي والإلكتروني، وبصفة عامة أصبح مورداً للمحتوى بكل أنواعه، ومسهماً قوياً في نشر المعرفة عن طريق المعلوماتية؛ مما أسهم في سد هذا الفراغ الضخم للمحتوى الإلكتروني العربي والإسلامي.

ويتفق هذا التوجه مع السعي الدائم والحرص الشديد -من قِبَل جهات عدة- على التواصل المعرفي والثقافي مع المواطن العربي بشتى الوسائل والطرق. ومن المعلوم أن الآلة التكنولوجية أصبحت وسيلة مناسبة للوصول إلى المواطن العربي لأهداف التعليم والتثقيف والتجارة وغيرها. كما أنها تعدّ -من جانب آخر- طريقة مناسبة لمدّ جسور التواصل الثقافي والحضاري بين العالم العربي والإسلامي وبقية دول العالم، وأن تكون واجهة عصرية تمثلنا ثقافياً وفكرياً أمام الآخر في عصر اختصرت فيه التكنولوجيا المسافات، وألغت الحدود، وحطمت الحواجز بين الدول.

وفي الوقت الذي دخلت فيه دور النشر العالمية منذ عشرين عاماً العالم الرقمي، ومعها دخلت كل هذه

محركات البحث مع خصائص اللغة العربية تنظيمياً ومعالجة واسترجاعاً وبحثاً، باستخدام أهم تقنيات معالجة النصوص العربية.

١ / اللغة العربية على الويب؛

تمثل مصادر المعلومات الرقمية باللغة العربية على الإنترنت قدرًا غير قليل من إجمال مصادر المعلومات الرقمية المتاحة للمستخدمين. ولا يمكن النظر إلى اللغة العربية لكونها لغة بعض مواد المعلومات المنشورة على الويب فحسب، وإنما تمثل اللغة العربية واحدة من أهم اللغات المستخدمة في استرجاع المعلومات وطرح بعض استراتيجيات البحث الخاصة بها والمميزة لها، مثل: آليات البحث بالمقابل أو المضادات ... وغير ذلك من خصائص البحث التي تتميز بها اللغة العربية عن غيرها من اللغات المتاحة على الإنترنت. وعلى ذلك فإن اللغة العربية يمكن تناولها من جوانب متعددة تخص عمليات النشر الإلكتروني والاسترجاع والمعالجة المعلوماتية لمواد المعلومات على قواعد بيانات الويب.

١ / ١ المحتوى العربي على الويب؛

يرى عادل محمد أحمد خليفة أن المحتوى الإلكتروني العربي من مواد المعلومات في البيئة الرقمية على غير المأمول لصنعه وتجهيز أرض خصبة لبناء المصادر الإلكترونية وتدعيم نشرها وتداولها. فلقد أكدت كثير من الدراسات على أن المحتوى المعلوماتي الرقمي العربي يقف على حدود فجوة رقمية بينه وبين المحتوى الرقمي الأجنبي المنشور إلكترونياً على الوسائط الصلبة أو على البيئة الشبكية. ويتجه الاقتصاد العالمي إلى ما

كما يشكل المحتوى فرصة قيمة للاستثمار، حيث تشكل عائدات شركات ألعاب الفيديو ما يصل إلى ٥٠ مليار دولار، علماً بأن من المتوقع أن تزيد عائدات هذه الصناعة بشكل كبير؛ مما دفع بكثير من الشركات الأمريكية العملاقة إلى استثمار مبالغ طائلة فيها .

كما أن للمحتوى قيمة تعليمية، حيث إن صناعة المحتوى هو وسيلة فعالة لبلوغ الغايات التعليمية الأربع التي نصت عليها منظمة اليونسكو، ألا وهي:

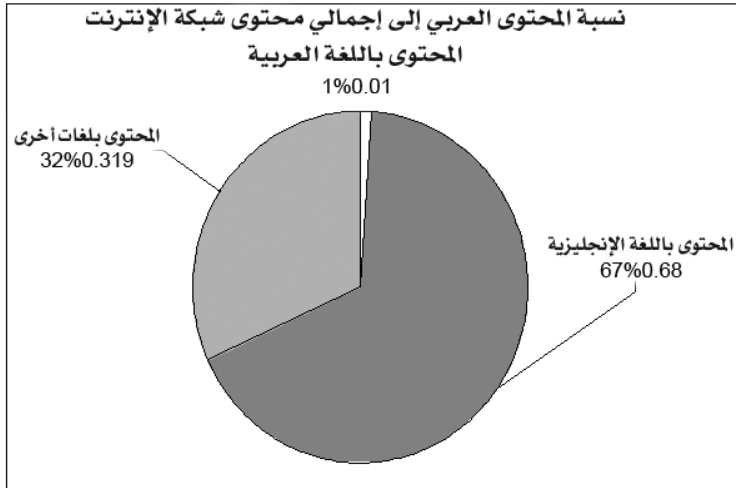
- تعلم لتعرف.
- تعلم لتكون.
- تعلم لتعمل.
- تعلم لتشارك الآخرين.

وأيضاً يشكل تطور المحتوى أحد أهم عناصر التطور التقني مستقبلاً، من حيث اندماج التلفزيون والإنترنت واستخدام خدمات الحزمة العريضة في نقل البرامج ونحوه.

الثقافات والحضارات عالم الحاسوب والإنترنت في صورة رقمية، إلا أن حجم التواجد الرقمي للثقافة العربية والإسلامية ما زال يشكل نسبة ضئيلة، ولا يتعدى ١٪ من حجم المحتوى العالمي على الإنترنت. إن هذه النسبة لا يمكن مقارنتها بنسبة تعداد سكان العالم الإسلامي (الذي تخطى عدده المليار نسمة)، ويشكل نحو ٢٥٪ من إجمالي سكان العالم.

ويوضح الشكل (١) أن المحتوى العربي المنشور إلكترونياً على شبكة الإنترنت يصل إلى المرتبة العاشرة في ترتيب المصادر المنشورة إلكترونياً باللغات الأخرى، وهو ما يمثل معه تراجعاً وتأخراً في صناعة المعلومات المنشورة إلكترونياً.

إن للمحتوى ذاته قيمة اقتصادية حيث كانت تقدر ينحو ١٧٨ مليار دولار خلال العام ٢٠٠١م حيث نمت بمعدل ٣٠٪ حتى وصلت إلى ٤٣٤ مليار دولار خلال العام ٢٠٠٦م^(٢).



الشكل رقم (١) المحتوى العربي مقارنة بمحتوى اللغات الأخرى

وكأنهما كلمتان مختلفتان تماماً، على الرغم من اتفاقهما في المعنى.

٢- قلة فاعلية المحركات العربية: حيث إن المحركات العربية تعمل تقريباً كواجهة عربية، ولكن النتائج توحى بأن المحرك يعمل تماماً كالمحرك الأجنبي.

١ / ٢ محركات البحث العربية:

من أكبر التحديات التي يواجهها الباحث العربي على وجه التحديد أن محركات البحث لا تبحث إلا عن الكلمات المطابقة مطابقة تامة لكلمات البحث. والسبب في ذلك أن مصممي محركات البحث يعتقدون أن عملية البحث هذه أنجح، وذات فائدة أكبر بالنسبة للمستخدم حين يكون البحث بالنص الحرفي فقط. مدعين أن توسيع مجال البحث من غير قيود قد يؤدي إلى نتائج بعيدة عن قصد المستخدم. لكن ذلك يعني أن على المستخدم نفسه التفكير بالأشكال المختلفة التي يمكن أن ترد فيها كلمات البحث وتجربتها كلاً على حدة.

وينبغي أن تستخدم عدداً مختلفاً من محركات البحث، وخاصة إذا أردت بحثاً شاملاً، إذ نادراً ما تجد النتائج نفسها في جميع محركات البحث. وهي كثيرة جداً، نذكر منها Alltheweb و msn، وهما محركا بحث عالميان يدعمان العربية بشكل جيد. أما محركات البحث العربية فيأتي في مقدمتها عجب Ajeeb، وهو أكثر محركات البحث العربية تطوراً، ويحوي بحثاً متقدماً، ويقدم خيارات اللواصق

لقد أحدث تطور المحتوى، وخاصة من خلال شبكة الإنترنت، تطوراً إعلامياً مهماً ليصبح الإعلام تواصلًا، لا مجرد بث من طرف لطرف على الرغم مما قد يثيره من مشاكل أمنية مثل الاستغلال السلبي للإعلام والإنترنت، من أجل نشر مبادئ هدامة أخلاقياً واجتماعياً لا بد للتصدي لها من خطاب عربي ذي محتوى سليم.

يبلغ عدد سكان العالم العربي نحو ٣٠٠ مليون نسمة، وهم يمثلون نحو ٥٪ من إجمالي سكان العالم، بينما تبلغ نسبة المحتوى العربي على شبكة الإنترنت نحو ٠,٠٠١٪ (واحد في الألف)، أي أنه يجب على المحتوى العربي أن يتضاعف نحو ٥٠ مرة لكي يصل إلى المستوى المطلوب. ناهيك عن ضحالة المحتوى العربي من حيث مكوناته وما به كثير من نقاط الضعف.

ويضاف إلى ذلك من خصائص المحتوى العربي أنه يواجه كثيراً من الصعوبات في عمليات البحث والاسترجاع، من خلال محركات بحث الويب، وذلك يرجع في وجوده الى عاملين هما :

١- صعوبة البحث بالمحركات الأجنبية: حيث إن المحركات الأجنبية مصممة خصيصاً للغتها الأم وليست للغة العربية. ومن ثم، نجد صعوبة في العثور على النتيجة المرجوة باللغة العربية. فمثلاً، إذا بحثنا عن كلمة «أحمد» بالهمزة وكلمة «احمد» بدون الهمزة يأتي المحرك الأجنبي بنتائج خاصة لكل منهما على حدة،

المجموعة الثانية :

اعتمدت هذه المجموعة من محركات البحث على تقنيات متقدمة لمعالجة اللغة العربية، ومن أبرز الأمثلة عليها: أراب فيستا، والإدريسي الذي أنتجته شركة صخر.

تتميز المجموعة الثانية بإمكانات إضافية مثل: البحث بالواصق، والبحث بالمشتقات . وقد استفاد محرك الإدريسي من التقنيات المتقدمة التي ابتكرتها شركة صخر في معالجة اللغة العربية، فتميز عن غيره بعدة أمور، أهمها:

١- البحث بالمترادفات: تستخدم للبحث عن نصوص عربية متشابهة المعنى كما يلي:

البحث عن كلمة مساعدة سيعطي من بين نتائج البحث كل الصفحات التي تحتوي كلمات عربية، مثل: مساعدة، ومؤازرة، ومعوونة ... إلخ.

٢- البحث بالمعاني والترجمة: تستخدم للبحث عن نصوص إنجليزية عن طريق كلمة عربية مجهل المستخدم معناها .

مثال: البحث عن كلمة «مساعدة» سيعطي من بين النتائج كل الصفحات التي تحوي كلمات إنجليزية، مثل help و assistance .

٣- إمكانية التعامل مع التشكيل بشكل جيد .

ومن الجدير بالذكر أن معظم محركات البحث العربية تقوم بعد معالجة الكلمة أو العبارة المراد البحث عنها بترجمتها إلى اللغة الإنجليزية، ليجري البحث عنها بعدة لغات في مواقع الويب المُفهرسة لديه.

والمشتقات والمترادفات والمتضادات والمعاني والبحث على مستوى الجذر. كما يعالج محرك البحث أراب فيستا ArabVista في موقع البحار مثل هذه المسائل العربية. والباحث المستقصي يلزمه أيضاً مراجعة محركات البحث والأدلة العربية الأخرى مثل Ayna و naseej، وفي كل منها نجد نتائج ربما لا نجدها في غيرها^(٣).

ظهر مؤخراً بعض محركات البحث التي تدعم البحث باللغة العربية، ويكمن السبب في قلة هذه المحركات وتأخر ظهورها إلى التقنيات المعقدة التي يحتاجها البحث باللغة العربية. إذ تختلف طبيعة اللغة العربية عن الإنجليزية، فاللغة العربية لغة صرفية (morphological)، بينما الإنجليزية لغة لصقية (affixational). ومن هنا كان لا بد للشركات التي تطرح محركات بحث عربية قوية أن تمتلك التقنيات اللازمة لمعالجة اللغة العربية آلياً.

وقد ظهر أثر ذلك في محركات البحث الموجودة التي انقسمت إلى مجموعتين^(٤):

المجموعة الأولى :

قلدت هذه المجموعة محركات البحث الإنجليزية؛ ولذلك فقد جاءت نتائجها ضعيفة لاعتمادها في البحث على المطابقة الحرفية لكلمات البحث، مما يتسبب في حجب كثير من المعلومات التي تتوافق مع الكلمات المراد البحث عنها (التي قد تختلف بأحرف زائدة بسيطة).

٢ / المعلومات العربية على Google:

بدأ محرك البحث Google عندما قام اثنان من طلاب الدكتوراة في جامعة ستانفورد بتأسيس جوجل سنة ١٩٩٨م. ويهدف إلى تقديم أفضل خدمة بحث على الإنترنت بجعل عالم المعلومات المتوافر على الشبكة في متناول الباحث . أما كلمة جوجل فتعني عدداً بملايين المليارات، ويعكس استخدام جوجل لهذه الكلمة إصرار الشركة على تنظيم الكمية الهائلة من المعلومات المتوافرة على الشبكة وفي العالم. حيث إنه محرك البحث الأول في العالم .. بمليارات الصفحات المفهرسة في قاعدة البيانات الخاصة به... وبسرعته الهائلة .. ومع ظهور منافسين جدد على الساحة ... يبقى جوجل في قمة محركات البحث.

يقدم جوجل أكبر محرك بحث في العالم بلا منازع؛ فعدد الصفحات التي يبحث جوجل فيها يزيد على ٨ مليارات صفحة (أما عدد الصفحات غير المفهرسة فيزيد عن؛ مليارات حتى أواخر ٢٠٠٤م)، ويوفر نتائج البحث لمستخدمين من كل أنحاء العالم، غالباً في أقل من نصف ثانية. واليوم، يلبي جوجل أكثر من ١٠٠ مليون عملية بحث في اليوم في مختلف لغات العالم مما يجعله مفضلاً لدى غالبية الباحثين^(٥). وهدفه ورسالته هما، تقديم أفضل خدمة بحث على الإنترنت؛ بجعل عالم المعلومات المتوافر على الشبكة في متناول الباحث ولفائده.



الشكل رقم (٢) واجهة استخدام محرك البحث (Google)

أمثل لما أصبح يعرف بلعبة الإنترنت التي أصبحت تدير رؤوس الكثيرين.

الحديث في هذا الموضوع كبير جداً (المواقع العربية وانتشارها) ومناسبة الحديث هذا في بحثنا هو عندما نبحث باللغة العربية هو أننا نبحث عن هذه المواقع.

لكن ما إحصائيات هذه المواقع والتواجد العربي على الإنترنت؟ «قال نائب رئيس فريق الأمم المتحدة لتكنولوجيا المعلومات طلال أبو غزالة إننا أصبحنا الآن أمام خيارين إما أن تصبح لغة الإنترنت هي الإنجليزية أو أن يتاح للغات العالم الدخول والتخاطب بأحرف لغتها ومنها لغتنا العربية؛ هذا إذا أردنا بناء مجتمع المعرفة العالمي. كما أن حجم مواقع الإنترنت العربية لم تتجاوز الصفر فعدد مستخدمي الإنترنت في العالم العربي لا يزيد عن ١٪ على حد قوله وهم المستخدمين الذين يمتلكون e-mail في الوقت ذاته، فلم يزد عدد المواقع العربية عن ٠,٠٠١ من إجمالي الموقع العالمية، فهناك مواقع عربية هزيلة تحتوي على صفحة أو صفحتين فقط في حين بعضها الآخر الذي يتجاوز آلاف الصفحات فحجمنا كدول عربية يصل إلى ٢٠٠ مليون نسمة، وعدد المستخدمين الدائمين قليلون بسبب غياب المحتوى العربي وعدم التمكن من استعمال الكمبيوتر والإنترنت ونحن في انتظار المستخدم الذكي العربي للإنترنت، الذي يستخدمها في أغراض التجارة والتعليم والتعامل

بدأ بداية قوية وهو الآن يعتبر محرك البحث المفضل لدى غالبية الباحثين المحترفين لديه واجهات تطبيق بلغات مختلفة من بينها العربية، وقد أثبتت الاختبارات كفاءة ممتازة لهذا المحرك عند البحث باللغة العربية ويوفر نتائج البحث للمستخدمين من كل أنحاء العالم بنحو ٣٥ لغة ، عموماً في أقل من نصف ثانية.

يتميز Google عن محركات بحث الجيل الأول المعروفة اليوم بتقنية إبداعية، فبدلاً من استعمال كلمة مفتاح أو تكنولوجيا البحث المتعدد metasearch، يعتمد Google على تقنية تصنيف الصفحة PageRank التي تضمن حلول أهم النتائج أولاً بشكل دائم^(١).

٢ / ١ قاعدة المعلومات العربية في Google:

عندما نتكلم عن البحث العربي يجدر التحدث عن التواجد العربي على الإنترنت من مواقع ومستخدمين فبالنسبة للمستخدمين العرب وصل عدد مستخدمي الإنترنت في العالم العربي إلى مليوني فرد، وهو رقم قد يكون ضخماً، إلا أنه مقارنة بالعالم رقم أقل من الطبيعي. واكب هذا العدد الكبير من مستخدمي الإنترنت عدد كبير أيضاً من المواقع العربية المتعددة الاتجاهات وذات الشخصيات المتعددة.. وأصبح فضاء الإنترنت مليئاً بالمواقع العربية الشخصية والمؤسسية الهادفة للربح وغير الهادفة للربح. بما يعكس تأثيراً على الثقافة العربية ينبغي رصده باستمرار وصولاً لاستخدام

ومن أخطاء المواقع العربية الاعتماد في تصميم الصفحات على نوع واحد من المتصفحات، مثل إكسبلورر أو نيتسكاب متناسين أن مستخدمي الإنترنت يتقاسمون استخدام هذين المتصفحين مناصفة، وأن كثيراً من المواصفات يدعمها أحد المتصفحين ويهملها الآخر. ويؤدي هذا إلى ظهور الصفحة بشكل سليم في أحد المتصفحات وظهور بعض الخلل فيها لدى الطرف الآخر.

ومن المشاكل التي واجهتها في البداية أنها لا تعرض بشكل جيد على المتصفحات، فلقد كانت تبدو عبارات غير مفهومة وعلامات استفهام ومربعات غامضة. وكان الحل في إيجاد متصفح عربي مثلاً سندباد، وثمة حل آخر حيث لجأ المصممون العرب إلى حيلة الصور، فحولوا نصوص المواقع إلى صور يمكن عرضها في كافة الحواسيب.. ولكن المشكلة كانت في ببطء التحميل. أما الآن ومع التطور الكبير الذي شهدته الإنترنت في العالم العربي فقد أصبحت المواقع تصمم بلغة html وهي لغة سهلة وواضحة جعلت من المواقع العربية في طليعة الركب الحضاري.

٢/٢ خصائص اللغة العربية على جوجل (Google):

عند الحديث عن محركات البحث العربية يجدر بنا التحدث عن خصائص اللغة العربية وهي ^(٩) :
١- اللغة العربية لغة اشتقاقية، أي أنها تعتمد على الجذر مثل الكلمات (دارس، دراسة،

اليومي، وهذه النسبة لن تتحسن إلا إذا فرضنا اللغة العربية كلفة تداول عبر الإنترنت - فمثلاً بدأ الصينيون يتوسعون في استخدام الإنترنت وأصبح النمو في استعمال الإنترنت باللغة الصينية أكبر من استعمالها باللغة الإنجليزية؛ لأن الصينيين يمثلون ثلث سكان العالم ولهم حجمهم عالمياً على العكس منا كدول عربية، لا يزيد تعدادها عن ٢٠٠ مليون نسمة فإمكانياتنا وحجمنا لا يدفعنا إلى الانعزال عن العالم الخارجي وعمل نظام إنترنت عربي خاص بنا. وهناك إحصائيات أخرى تقول إنه يوجد ١٤ مليون مستخدم للإنترنت من أصل ٢٠٠ مليون عربي ^(٧) .

أما عن تصميم المواقع العربية ومشاكلها مع المتصفحات فله عدد من السمات الأساسية التي تميزها ومنها أن تصميم الموقع بلغة غير العربية، وهو من الأخطاء التصميمية المنتشرة في كثير من المواقع العربية، وهو أول مؤثر تغريبي يصيب العقل العربي بثقافة أجنبية، فالقصد من إنشاء صفحة على الإنترنت هو إيصال فكرة أو معلومة أو ترويج منتج معين لأكبر مجموعة من الناس، ويبدو أن هذا الأمر يغيب عن كثير من أصحاب المواقع العربية، الذين يعتبرون وجود صفحات بلغة غير عربية أمراً «راقياً»!! أو أنهم يفتقرون إلى المعرفة التي تسمح لهم بإنشاء صفحات عربية؛ بسبب عدم إلمامهم بلغتهم، أو لتكليفهم جهة لا تقدم صفحات عربية مهمة إنشاء الموقع ^(٨) .

كلمة "طالب" من المفترض أن يتم الحصول على الكلمات "طلبة" و"طلاب" و"طالبات" وغيرها من الكلمة المشتقة من فعل "طلب" وهذا بالطبع من الصعب برمجته في محرك البحث؛ لأنه يتطلب إرجاع الكلمة إلى جذرها اللغوي، ومن ثم اشتقاق جميع الكلمات منه، ومن ثم البحث عن هذه الكلمات المشتقة. أما اللغة الإنجليزية فهي لغة إلحاقية أو لصقية حيث إن جميع مشتقات الكلمة عبارة عن إلحاق حروف معينة بها مثل "er" و"ed" وغيرها من الملحقات الأخرى (مع وجود بعض الاستثناءات البسيطة).

أيضاً فإن محركات البحث الحالية لا يمكنها أن تفهم اللغة العربية كما تفهم اللغة الإنجليزية، مثلاً عند البحث بكلمة عربية فيها خطأ إملائي لا يصحها محرك البحث بل يأخذها كما هي، بينما إذا أخطأت في كتابة كلمة إنجليزية فإن محرك البحث يحاول تصحيحها فمثلاً ابحث عن كلمة linux ستجد أن جوجل يصحح لك الخطأ الإملائي ويقترح عليك الكلمة الصحيحة وهي Linux، لماذا لا يصحح لنا جوجل أو أي محرك بحث أخطاءنا الإملائية؟

وأيضاً لا تستطيع محركات البحث التفريق بين الهمزات، أو لا تستطيع فهم أن كلمة أحمد هي نفسها احمد أو إحمد، لذلك هي تقدم نتائج مختلفة ولا تستطيع اقتراح الكلمة الصحيحة. وهناك مشكلة أخرى وهي (ال) التعريف، فكثير من

- ١- مدرسة....)، التي تعود إلى الجذر (درس).
- ٢- تكتب من اليمين إلى اليسار.
- ٣- لا بد من اتصال الحروف بعضها ببعض لتكوين الكلمة.
- ٤- الحرف الواحد له أكثر من شكل حسب موقعه من الكلمة.
- ٥- وجود علامات الضبط بالشكل (فتحة، ضمة، كسرة، سكون، تنوين).
- ٦- كثرة المترادفات في اللغة العربية (مساعدة، دعم).
- ٧- هناك حروف قد تكتب بأكثر من شكل في الموقع نفسه.
- ٨- إمكانية التراكم الرأسي.
- ٩- عرض بعض الحروف أكبر من الآخر.
- ١٠- الكلمة في اللغة العربية كلمة مركبة، حيث يدخل ضمن الكلمة الجنس، وصفة المخاطب، والعدد، والزمان، وغيرها.
- ١١- هناك كلمات بها حروف تنطق ولا تكتب مثل الرحمن.

هذه بعض الخصائص في اللغة العربية أغلبها يؤثر على عملية البحث، وبعضها لا يؤثر.

البحث باللغة العربية أصعب من البحث باللغة الإنجليزية؛ لأن اللغة العربية لغة اشتقاقية بينما اللغة الإنجليزية لغة لصقية؛ فمثلاً عند البحث عن

٣- تجاهل الحركات والتنوين والحروف الغريبة مثل حرف السمد والنجوم وغيرها: «^&*#@!»،

٤- تجاهل (ال) التعريف، فكلمة العلماء يجب أن تكون هي نفسها كلمة علماء لدى محرك البحث.

٥- تجاهل الكلمات التي تستخدم كثيراً مثل: من، وأين، وحرف الواو، حينما يسبق الكلمة)، مثال: وقال لي فلان كلمة قال هنا يجب أن يتجاهل محرك البحث حرف الواو الذي يسبقها فيتعامل مع كلمة (وقال) مثل ما يتعامل مع كلمة (قال).

أعتقد أن أكثر الحلول تعتمد على تجاهل الحروف أو معاملتها بالمثل، فكما قلت من قبل يجب أن يفهم محرك البحث أن (أ) هو نفسه (ا) أو (إ)، لكن الباحث قد يخطئ، وللمزيد من الدقة لا بد من دراسة النحو والصرف وفنون اللغة الأخرى لإنشاء محرك بحث يفهم العربية بشكل كبير.

وأعتقد أن شركة صخر قادرة على فعل شيء في هذا المجال، فليدهما أبحاث تتعلق باللغة وقامت بإنتاج كثير من المنتجات والخدمات التي تقدمها الآن لشركات عالمية مثل IBM وإنتل، لكن نحن كمجموعة من الأفراد علينا ألا نعول على شركة، يجب أن نعمل بأنفسنا.

كلماتنا العربية تبدأ بحرفي التعريف، ومن الخطأ تصنيف كلمة الحضارة تحت الحرف ألف ثم لام ثم هاء، بل يجب تجاهل حرفي التعريف وتصنيف الكلمة مباشرة تحت حرف الهاء، لكن محركات البحث لا تفهم هذه القاعدة.

وأيضاً يتجاهل جوجل الكلمات الشائعة في الإنجليزية مثل the و on و how وغيرها، ولدينا في العربية كلمات مماثلة مثل: في، ومن، وأين، لكنه لا يتجاهلها لأنه لا يفهم العربية. وكذلك محركات البحث الأخرى^(١٠).

هذه مجموعة من المشاكل التي أراها في محركات البحث، لو تم حلها بنجاح يمكننا أن ننتقل إلى مرحلة أكبر وأكثر تعقيداً، وهي تتعلق بالنحو والإملاء والصرف في اللغة العربية، ويعني ذلك أننا استطعنا حل جزء كبير من مشكلة محركات البحث مع لغتنا.

هناك حلول مقترحة وهي:

١- تجاهل حرف (أ) واعتبار أن (أ) هو نفسه (ا) أو (إ).

٢- إنشاء قاموس إملائي يقترح الكلمة الصحيحة على الباحث في حال أخطأ الباحث في الإملاء، ويجب أن يعطي الباحث أقرب كلمة لما يبحث عنه، فلو بحث عن كلمة (علمئ) ف عليه أن يقترح عليه كلمة صحيحة وهي: علماء.

٣/ مشكلات البحث باللغة العربية:

تتسم اللغة العربية بعدد من السمات التي تتفرد بها عن غيرها من اللغات الأخرى، وعلى ذلك فإن ثمة تأثير يمكن أن تحدثه اللغة داخل نظام الاسترجاع، ويجلو أثره أكثر على التحقيق والاستدعاء داخل نظام الاسترجاع باللغة العربية. ويختص هذا الجزء من الدراسة باستيضاح خصائص اللغة العربية وكيفية تطويعها بشكل يحقق الكفاءة لنظام استرجاع المعلومات، يعتمد في بنيته على اللغة العربية. والكيفية التي تكون عليها مصطلحات اللغة العربية داعمة مقياسي الاستدعاء والتحقيق، بما يرجوه المستفيد من نظام الاسترجاع.

١/٣ مشكلات البحث الحر باللغة العربية في**البيئة الرقمية:**

ينطوي مصطلح البحث الحر على استخدام البحث باللغة الطبيعية داخل النصوص الكاملة لمصادر المعلومات. وينتج عن العمل باللغة الطبيعية في بحث المصادر العربية عدد المشكلات التي تعكس خصائص اللغة العربية، ويمكن حصرها في الفئات التالية:

١/١/٣ الاسترجاع اللفظي:

وهي مشكلات قد حلها وجود برنامج التحليل الصرفي المضاف إلى معظم محركات بحث، ولكنها ستظل قائمة لو لم يزود محرك البحث ببرنامج التحليل الصرفي. ومن ثم ينوه عنها الباحث هنا للتعريف بأنواع هذه المشكلات المرتبطة بخصائص اللغة العربية في حالة غياب المحلل الصرفي، وهي:

أولاً: صيغة الفعل:

وهي الصيغ المختلفة لمشتقات الفعل، فالبحث عن كلمة (الحج) مع تحديد خيار البحث بالمشتقات، يستدعي وحدات تتضمن (الحج، حجة، حاج، يحجج... إلخ).

ثانياً: الإملاء ورسم الكلمات:

وهي المشكلات الناجمة عن الأشكال المختلفة لكتابة الكلمات، خاصة في حروف الهاء الختامية للكلمة مع التاء المربوطة، والياء اللازمة والألف المقصورة المرسومة بالياء، وهمزة القطع والوصل. وعلى ذلك فإن بحثاً لكلمة (احمد) سوف يستدعي أيضاً الشكل الآخر للكلمة (أحمد).

ثالثاً: الحروف والأدوات «قائمة التوقف»:

وهي التمييز بين الحروف والأدوات التي قد تعوق البحث، مثل التمييز بين (ال) التعريف كواصق في الكلمة أو كأصل، فكلمة مثل «الله» تعد فيها (ال) أصلية، بينما كلمة مثل (الكتاب) لا تعد كذلك. وتعمل نظم الاسترجاع على حذف (ال) التعريف عادة في مقدمة الكلمات، ولكن ذلك يتطلب وجود قائمة كلمات التوقف التي تستبعد بعض الكلمات، ومن ثم يكون من المناسب هنا وضع الكلمات التي تأتي فيها (ال) أصلية، مثل «الذي» و«التي» وغيرها.

٢/١/٣ الاسترجاع بالمعنى:

وهي المشكلات التي تتعلق بالكلمات المفتاحية المستخدمة في وصف المعلومات الرقمية، ومن ثم فهي تتعدى الحدود اللفظية للمعنى الضمني. ومثل

الترادف على تشتت المداخل في الكشف، إلا أن البيئة الرقمية لنظم الاسترجاع أعادت صياغة مشكلات اللغة العربية بمستوى بحث آخر، خلاف ما يكون في البيئة التقليدية. وقد كانت الإحالات هي الحل الأمثل لمثل هذا النوع من مشكلات الاسترجاع اليدوي لمعلومات اللغة العربية^(١٣).

ويمكن تصوير الترادف في مثاله الشهير عن الكلمات الثلاثة المستخدمة لوسيلة الاتصال بالهاتف، وهو (الهاتف، والمسرة، والتليفون). إن بحثاً في محركات البحث التقليدية من شأنه أن يتسبب في ضياع جزء من وحدات المعلومات المنتمة للموضوع.

ب- الاشتراك اللفظي:

يقصد بالألفاظ المشتركة أو المشترك اللفظي مجموعة من المعاني والمفاهيم المختلفة التي تتوحد في مجموعة حروف واحدة مكونة كلمة ذات شكل واحد، بهذا فإن الباحث عند استخدامه لأحد الكلمات المتسمة بهذه السمة سوف يكون أمام فيض من المعلومات التي لا حاجة له بها، ومن أمثلة ذلك أن يقوم باحث بالبحث عن معلومات عن (قطر) فيدخل كلمة (قطر) فتأتي أشكال رياضية تمثل (قطر) الدائرة. ولقد أثمر جهد نظم الاسترجاع بالاعتماد على تقنية التحليل المعجمي آلية البحث بالتشكيل، وعلى الباحث عند إدخال كلمة البحث ثم تحديد خيار البحث بالتشكيل لتعيين المعنى المطلوب اعتماداً على شكل الكلمة. إلا أن قواعد البيانات

هذه المشكلات تعمل بعض محركات البحث على إيجاد حلول لها، إلا أنها غير قاطعة، ومن أهم هذه المشكلات:

أ- المترادفات:

يعرف الترادف على أنه: «توالي الألفاظ المفردة الدالة على مسمى واحد باعتبار معنى واحد فيخرج عن هذا دلالة اللفظين على مسمى واحد لا باعتبار واحد، بل باعتبار صفتين كالصارم والمهند، أو باعتبار الصفة وصفة الصفة كالفصيح والناطق»^(١١). وأضاف عبد الرحمن المحسني إلى ذلك أن ثمة قضية لغوية تتعلق بالترادف، وهي إنكار الترادف وإثباته، فقد انقسم أهل اللغة إلى فريقين في هذا الشأن، ويرى فريق الإنكار أن الشارع حكيم ولا توجد كلمة إلا ولها دلالة خاصة بها، وتبعاً لذلك فلا ترادف في اللغة. أما فريق الإثبات فيرى أن الترادف موجود باللغة، وأن هناك كلمات عربية كثيرة مترادفة لم تأت عبثاً وإنما هناك مقاصد محددة لذلك، وأكد هؤلاء رأيهم بالأدلة العقلية والآيات القرآنية الكريمة^(١٢).

أما من منظور نظم استرجاع المعلومات، فإن الترادف يسهم في ضياع جزء من التحقيق للمستفيد في استرجاع المعلومات، على عكس ما يكون الأمر عليه في مشكلة الاشتراك اللفظي، حيث يعاني المستفيد من ارتفاع نسبة التشويش (الشوشرة) في النتائج المسترجعة. وعلى الرغم من أن المشكلة لدى الصوابع توقفت عند الكشافات اليدوية وتأثير

بالتركيز على محرك البحث Google.

١/٤ آليات البحث باللغة العربية على جوجل (Google):

تختلف اللغة العربية عن غيرها من اللغات الأخرى في عمليات تنظيم المعلومات في درجة المرونة المطلوبة لتنظيم المعلومات واسترجاعها، ومن ثم فإن آليات بحث المعلومات داخل اللغة العربية تختلف عن تلك المستخدمة في اللغات الأخرى مثل اللغات اللاتينية بأنواعها المختلفة. ودخل استرجاع اللغة العربية عهداً جديداً بالاعتماد على تقنيات التحليل الصرفي والمعجمي لألفاظ اللغة العربية وتراكيبها، وانتفت مع هذه التقنيات معظم مشكلات الاسترجاع الآلي للمعلومات في اللغة العربية. وقد اعتمدت بعض محركات البحث العامة باللغة العربية على مثل تلك التقنيات، ومن أهمها محرك عجيب Ajeeb الذي انعكس استخدام تقنيات التحليل الصرفي به في آليات البحث بالمعنى والمشتقات والجذر والمترادفات.

وسيتم الاعتماد على محرك البحث Google لدراسة خصائص استرجاع المعلومات في بيئة اللغة العربية. ويوضح الجدول رقم (١) آليات البحث في اللغة العربية وتطبيقها في محرك البحث Google على الويب:

التي تعنى بالكلمات لا تقوم بتشكيل الكلمات المفتاحية، ولم يظهر في الأفق حتى الآن التقنيات التي تسهم في إصلاح عمليات البحث في مصادر اللغة العربية.

٣/١/٣ اللغة والثقافة:

تشكل الثقافة والبيئة التي يتم تكشيف المصادر المعلوماتية بها عاملاً مهماً في اختيار الكلمات المفتاحية عن محتوى هذه المصادر. ويمكن لمصدر واحد أن يستخدم لإيصال أكثر من معنى، وقد لا تكون هذه المعاني في حالة تباين فقط، وإنما في حالة تضاد في المعنى؛ على سبيل المثال فإن صورة واحدة لطفل فلسطيني يحمل حجراً يرمي به جندياً صهيونياً، يمكن استخدامها لإيصال معنى النضال، كما تستخدم على الجانب الآخر في الإعلام الغربي لإيصال معنى الإرهاب، وهو البعد الثالث في التعامل مع تحليل مصادر المعلومات متعددة الوسائط.

٤ / آليات البحث باللغة العربية على جوجل (Google):

يرسم واقع البحث باللغة العربية على محركات البحث ملامح المشكلة الحقيقية التي تواجهها مصادر المعلومات الرقمية باللغة العربية داخل قواعد محركات البحث، حيث تكشف دراسة تطبيقية بسيطة كالتي قام بها الباحث هنان أن هناك قصوراً شديداً في معالجة المعلومات العربية الرقمية، من حيث بنية اللغة العربية وطبيعة المحارف العربية، وقدرات البحث داخل محركات بحث الويب، وهنا

الجدول رقم (١) آليات بحث اللغة العربية في محرك البحث Google

حالة آلية البحث	آليات بحث اللغة العربية
مستخدمة	البحث بالتطابق
غير مستخدمة	البحث غير المطابق باللواصق
غير مستخدمة	البحث بجذر الكلمة
غير مستخدمة	البحث بالتشكيل
مستخدمة مع رموز ٩ ×	البحث الحر
غير مستخدمة	البحث بالترادف
غير مستخدمة	البحث بالمتضادات
غير مستخدمة	البحث بالمعاني
مستخدمة	البحث بالترجمة
مستخدمة	التعرف إلى الأخطاء العربية الشائعة

من خلال المصحح الآلي، الذي يحلل النصوص المدخلة إملائيًا.

ثالثًا: اختلاف كتابة بعض الحروف بين البلدان العربية: فمثلاً يكتب المصريون الياء بدون نقط، أما في الدول الأخرى فإنها تكتب بنقط. عدم التنبه إلى هذه النقطة قد يؤدي إلى عدم عرض النتائج كاملة أثناء البحث، وهو ما يحصل في Google.

رابعًا: يستخدم محرك البحث Google بعض ملامح البحث التي تتسحب على استخدام وبحث مختلف اللغات مثل البحث بالتطابق والبحث الحر باستخدام رموز الامتدادات المختلفة. وبالإضافة إلى ذلك فقد لاحظ الباحث عدم تخصيص اللغة العربية بمجموعة أو واحدة على الأقل من آليات البحث التي تخدم البحث باللغة العربية، وموافقة

يلاحظ الباحث من الجدول السابق مجموعة من الإشارات تتمثل فيما يلي :

أولاً : خصائص اللغة العربية وخصوصياتها: ومن ذلك اللواصق الموجودة في اللغة العربية من السوابق واللواحق. على عكس Google فإن محرك البحث الإدريسي اعتنى بهذه النقطة، ومن ثم فإن النتائج التي لا تظهر في Google نتيجة لعدم الانتباه لهذا الموضوع ستكون ظاهرة في محرك البحث الإدريسي.

ثانيًا: عدم أخذ الأخطاء الإملائية بعين الاعتبار: حيث إن الأخطاء الإملائية بشكل عام والمتعلقة بالهمزة بشكل خاص كثيرًا ما تشكل مشكلة عند الشروع بالبحث. بينما لم تلتفت جوجل Google لهذه المشكلة فإن الإدريسي قد تلافاها

١/٢/٤ تقنيات اللغة العربية:

يذكر محمد سالم غنيم أن هناك أهمية كبرى لتطويع التقنية لخدمة اللغة العربية. فمع ظهور تقنية المعلومات، وعلى رأسها الحاسبات الآلية، برزت مشكلة تطويع هذه التقنية لخدمة اللغة العربية؛ حيث إن تصميم تلك الحاسبات تم لخدمة المجتمعات التي أنتجتها، فواجهت محاولات استخدام الحاسب الآلي في البلاد العربية مشكلة تطويعه لمعالجة النصوص العربية، إذ بات على البلدان العربية مواكبة هذا التيار الحضاري، والانتفاع بالمبتكرات الحديثة؛ خدمة لأغراضها الوطنية والقومية، وخاصة أننا نشهد اليوم ثورة ونهضة علمية في المنطقة العربية، بعد أن واكبت الحاسبات الإنسان في نشاطاته كافة.

ويمكن وصف تلك الإشكالية التي تمر بها نظم المعلومات العربية في: «أنه ينبغي ألا يغيب عن بالنا أن اللغة العربية من أهم - إن لم تكن أهم- العوامل المحددة لهوية النظام العربي للمعلومات والمؤثرة فيه، ويتوقف تطوير هذا النظام، والارتقاء بمستوى فعاليته وكفاءته وقدرته على استثمار هذه اللغة بكامل طاقتها في جميع عناصر النظام، بدءاً باستعمالها في قنوات النشر على اختلاف فئاته ومستوياته، وانتهاءً بتطوير الحاسبات الإلكترونية وغيرها من تقنيات المعلومات القادرة على التعامل مع العربية بشكل فعال، مروراً بالقوائم الاستنادية ولغات التكشيف، وغير ذلك من الأدوات المتأثرة بشكل ما بالنظام اللغوي العربي. ومما لا يخفى على

خصائص البحث بالعربية دون غيرها من اللغات الأخرى.

خامساً: تحتاج اللغة العربية إلى مجموعة من آليات البحث كتلك المشار إليها في معالجة اللغة الطبيعية في النصوص العربية، ولم تزل محركات بحث النصوص العربية تعمل بآليات بحث عامة تتوافق مع الكلمات المفتاحية في مختلف اللغات الأخرى؛ مثل آليات المنطق البولييني، والبحث الحر، والتطابق، والبحث بالتقارب... وغيرها من الآليات الأخرى. ولقد عملت المحركات العربية حديثاً على استخدام التقنيات المساعدة في بحث اللغة العربية مثل تقنيات التحليل المعجمي والتحليل الصرفي، ولا يمكن بحث النصوص العربية المستقلة أو التابعة لوسائط الصور والصوت دون الاعتماد على التقنيات السابقة. لقد خلت محركات بحث العينة السابقة حتى المستخدم للعربية منها من آليات بحث تعمل خصيصاً لبحث وحل مشكلات استرجاع المعلومات من النصوص العربية، لخلوها من تقنيات التحليل الصرفي والتحليل المعجمي. واعتماداً على التقنيات السابقة يمكن إضافة مجموعة من آليات البحث إلى محركات بحث اللغة العربية.

٢ / ٤ متطلبات البحث باللغة العربية:

تتطوي معالجة اللغة العربية في بيئة المعلومات الرقمية على الإنترنت على مجموعة من المتطلبات تتمثل أهمها في معالجة وتقنيات التعامل مع اللغة العربية الرقمية، أيضاً التوصل إلى برمجيات تختص بحل مشكلات اللغة العربية الناتجة عن ثراء موسوعية هذه اللغة.

١- البحث بالتطابق:

يطلق على هذا النوع من البحث عدة مسميات منها البحث بالكلمة أو البحث بالمطابقة. وتعتمد هذه الطريقة على وضع كلمة محددة يبحث بها نظام الاسترجاع داخل محتوياته، على أن المضاهة هنا اعتماداً على الشكل نفسه الذي أقره المستفيد لإجراء البحث؛ فإذا ما استخدم المستفيد مثلاً كلمة (عربي) فإن النظام يعمل على استرجاع كل الوحدات التي تحتوي على كلمة عربي، بينما لا يسترجع النظام وحدات الكلمات الأخرى مثل (عرب، عربية، عربية). ويحتاج المستفيد إلى هذا النوع من البحث في حالة طلب معلومات عن موضوع له مخصص، بحيث يهتم المستفيد بالتحقيق أكثر من الاستدعاء الذي ستكون نسبته في هذه الحالة قليلة جداً.

٢- البحث غير المطابق «البحث باللواصق»:

يختلف هذا الشكل من البحث عن الشكل السابق في توسيع نطاق البحث داخل النظام. والبحث غير المطابق للكلمة مؤداه أن يسترجع النظام مختلف الأشكال للكلمة الواحدة، مع استدعاء أشكال الكلمة مع كل اللواصق التي تسبقها أو تليها في وحدات نظام الاسترجاع. فمثلاً إذا أراد الباحث البحث عن كلمة (عرب) فإنه سيحصل على (عرب ، عربي، عربية، عربيات) وكذلك يسترجع جميع الكلمات التي تحمل سوابق مثل (العربي، وللعربي، وبالعربي) ومثلها للكلمات الأخرى^(١٦). ويحتاج المستفيد لمثل هذا الشكل من البحث إذا ما أراد الحصول على كل وحدات المعلومات حول موضوع محدد يتمثل في كلمة محددة.

أحد أن تقنيات النظام العربي للمعلومات تعتمد على فئات موائد التطورات التقنية في المجتمعات المتقدمة، إن لم تكن نفايات هذه التطورات^(١٤).

يعد التحليل الصرفي لب عمليات معالجة اللغة الطبيعية، حيث تقوم نظم الاسترجاع بتوفير هذه الخاصية لإنتاج كثير من إمكانيات البحث المتولدة عن استخدام التحليل الصرفي لكلمات اللغة العربية. تبدأ هذه العملية بإضافة المعاجم اللغوية الآلية لنظام الاسترجاع، ولا ينسحب التحليل الصرفي على كل كلمات اللغة العربية وإنما يستثنى من ذلك الحروف وأدوات التركيب ... وغيرها من الكلمات التي لا تدل بمفردها على معنى معين. ويعد الإدريسي <http://www.alidrisi.com/>، وهو محرك بحث أنتجته شركة صخر للحاسبات، من أهم نظم استرجاع المعلومات التي تعمل بتقنية التحليل الصرفي، مما أدى إلى بزوغ إمكانيات بحث متقدمة لا ترى في غيره من محركات البحث العربية. لقد بات التحليل الصرفي الحل لمعالجة بحث اللغة العربية، بعد تناوله من جانب الصونين وغيره، يعتبر محركات بحث المعلومات العربية تربة خصبة تنمو بها تطبيقات التحليل الصرفي، حيث إن تحليل موضوعات مصادر المعلومات باستخدام اللغة الطبيعية يحقق الهدف المرجو في استرجاع المعلومات.

وتتمثل تقنيات البحث المتقدمة المتولدة عن التحليل الصرفي للغة العربية فيما يلي^(١٥):

الاختلاط بين الأشكال المختلفة للكلمة الواحدة، وكان الحل البديل هو وضع كلمة أخرى بجانب الكلمة المختلطة للاسترشاد بمعنى الكلمة الأخرى. ويعطي هذا الشكل للمستفيد أقصى درجات الدقة في الحصول على التحقيق المطلوب. وتتمثل طريقة استخدام شكل البحث في مجموعة بدائل لشكل الإعجام للكلمة يعرضها نظام الاسترجاع من واقع التحليل الصرفي للغة داخل النظام.

٥- البحث بالمقاطع (أجزاء الكلمة):

يعطي هذا الشكل مجموعة من المعاملات أو الحروف التي يستخدمها الباحث في تشكيل استراتيجية بحثه كما يريد، وتتركز هذه الحروف في علامات (٩، ×)، على أن استخدام هذه العلامات يختلف بين نظام وآخر، ويوجد جانبان من استخدام هذه الآلية في البحث هما استخدام معاملات البحث بالمقاطع في منتصف الكلمات، أو استخدام المعاملات في نهاية الكلمات؛ أما الجانب الأول من الاستخدام وهو في منتصف الكلمات فيأتي إما لعدم معرفة الباحث بالحروف الصحيحة المكونة لكلمات البحث، مثل البحث عن (عرب) تسترجع كلمات عرب، عراب، وأما الجانب الثاني من استخدام معاملات البحث الحر فيأتي في نهاية الكلمة ويستخدم للتخلي عن البحث بكلمات متعددة تشترك في أصل واحد مختلف الامتدادات وتستخدم العلامة (×) للتعبير عن مجموعة من الحروف التي تأتي في نهاية الكلمة مثل (عرب×) فإنها تسترجع "عربي، عربية، عربيات... إلخ، مع

٣- البحث بجذر الكلمة «حل مشكلة الاشتقاق»:

يأتي هذا الشكل في المستوى الثالث من توسيع نطاق البحث، وهو أهم أمثلة استخدام التحليل الصرفي للغة العربية، ويعني إجراء تحليل الكلمة إلى الحروف الأصلية بها واستخدامها نواة للبحث داخل نظام الاسترجاع. ومثال على ذلك فإن بحثاً عن معلومات تتعلق بـ (عرب) سوف تؤدي إلى استرجاع وحدات من داخل النظام تتضمن كلمات (عرب، عربي، عربية، عروبة... إلخ) (١٧) ويؤدي هذا الشكل من البحث بطبيعة عمله إلى استدعاء عدد غير محدود من مصادر المعلومات، عملاً على تجميع كل ما يتعلق بالجذر موضع البحث. ويترتب على نوع البحث نوع المستفيدين، فالمستخدم لهذا الشكل من البحث يلجأ إليه عند الحاجة إلى كل ما يتعلق بالموضوع سواء من قريب أو بعيد، لما يمكن استدعاؤه من معلومات. ولكن يعيبه أنه قد يأتي بكلمات ليس لها علاقة بالموضع مثل (عراب، إعراب).

٤- البحث بالتشكيل لكلمات المشترك اللفظي:

يطلق على البحث بالتشكيل مسمى «فك اللبس» كما ورد في بعض مصادر الحديث عن أشكال البحث في اللغة العربية. ويخدم هذا الشكل المستفيدين في تفادي أحد أهم عوائق الاسترجاع في اللغة العربية وهو المشترك اللفظي للكلمات، مثل (علم، علم، علم). وقد عانى المستفيدون كثرة

في إجراء البحث على المعجم الآلي، حيث تحديد المترادفات لكلمة البحث، واستخدامها في استدعاء مصادر المعلومات من وحدات النظام التي تحتوي على مرادفات كلمة البحث. فمثلاً يؤدي البحث عن كلمة «تجهيز» إلى استرجاع جميع المستندات المفهرسة، التي تحتوي على كلمة (تجهيز) أو إحدى مترادفات مثل: إعداد، وتهيئة... وغيرها من مرادفات الكلمة إن وجدت. ويمثل هذا النوع من البحث أهمية بالغة في البحث عن مصادر المعلومات العربية. ويعد محرك البحث الإدرسي^(١٩) من أفضل أدوات البحث العربية التي تعمل باستخدام تقنية التحليل الصرفي، بما يمكن معه إنتاج آليات بحث مثل البحث بالمترادفات.

٨- البحث بالمتضادات:

يستخدم المستفيدون هذا الشكل في البحث من أجل استرجاع مجموعة الوحدات التي تتعلق بالكلمات ومتضاداتها. وعلى ذلك فإذا ما رغب المستفيد في البحث عن كلمة (السلام)، ويختار البحث بالمتضادات فإن نظام الاسترجاع يستدعي أيضاً كلمة (الحرب) ... وغيرها. وهذه الطريقة ربما تكون مفيدة في البحث في النصوص، حيث تأتي كلمات مرتبطة ببعضها مثل (الحرب والسلام)، (الأبيض والأسود)، (الخير والشر) و(الجنة والنار)، ولكن بالنسبة لمواد المعلومات المصورة فإن هذا ربما يستدعي صوراً مخالفة لما هو مطلوب. وأفضل أدوات البحث العربية في ذلك محرك البحث الإدرسي.

التأكيد على أنها لن تسترجع كلمات مثل «العربي»، وبالعربي... إلخ»^(١٨).

٦- البحث بحساسية الحالة:

يعبر هذا الشكل عن المصطلح case sensitive، ويطلق عليه في بحث الإدرسي مراعاة حالة الأحرف، ويستخدم المستفيد شكل البحث في حالة ما إذا أدت الأشكال المختلفة لحروف الكلمة معاني مختلفة، يعمل البحث بحساسية الحروف أو شكل الحروف على البحث عن اتجاهات موضوعية مختلفة لموضوع واحد ولكلمة واحدة؛ حيث يعمل شكل الكلمة على بيان جانب من الموضوع، مثل كلمة «إعلام» فإن استخدام البحث بدون حساسية الحروف يسترجع كلمات أخرى مثل «إعلام». حيث يعامل حرف الألف كحرف واحد دون الأخذ في الاعتبار حالة الهمزة (ا، أ، إ، آ) وتستخدم هذه الآلية في محركات البحث متعددة الموضوعات، إلى جانب اختيار شكل الحروف كواحدة من أدوات ضبط المصطلحات في عمليات الكشف بما يصنع معه تحديداً أكثر للمصطلحات الكشفية. ولم يتعامل أي من محركات البحث السابقة مع البحث بشكل الحروف، وهذا يرجع في رأي الباحث إلى أن مثل هذه الآلية تحتاج إلى مرحلة متقدمة من البحث من جانب المستفيدين، الأمر الذي لا يتناسب مع الاستخدام العام للويب.

٧- البحث بالمترادفات:

في هذا النوع من البحث يركز نظام الاسترجاع

٩- البحث بالمعاني:

المعلومات العربية داخل محركات البحث وتطويره. وقد ذكر عبد الغني أبو العزم أن شركة صخر شرعت منذ بداية الثمانينات من الألفية الثانية في تحضير قاعدة مشروع معلوماتي على صعيد العالم العربي، وكانت تملك تصورًا يحكمه الوعي بمدى السرعة الهائلة التي يعرفها قطاع الإعلام والتواصل في مجال التقنية الإلكترونية وما تحقق من منجزات أوروبية وأمريكية في هذا الميدان في ضوء التجارب الأولى التي أنجزت، وما أصبحت تتوفر عليه من بنوك المعطيات وحواשב ضخمة لتخزين المعلومات والمفردات بكميات هائلة جدًا^(٢٠).

لقد تمكنت شركة صخر منذ شروعها في وضع برامجها في خانة التطور التي تعرفه الصناعة الإلكترونية دوليًا وعلى قواعد وكفاءات علمية، وذلك لجعل اللغة العربية في مصاف اللغات العالمية، واستطاعت أن تحقق خلال عقدين من الزمن طفرة هائلة، وفجرت طاقات الإبداع والإنجازات في مجال تطوير أعمال وبرامج هي الآن تشكل خزانة مهمة في المدارس والمؤسسات والمعاهد والإدارات. وكانت أهم البرمجيات الناتجة عن جهود شركة صخر تتمثل في البرمجيات الآتية^(٢١):

أولاً: المحلل الصرفي:

يعد المحلل الصرفي الآلي من الآليات الأساسية للتعامل مع طبيعة الكلمات، سواء كانت مجردة أو مرتبطة بزوائد ولواحق، وذلك باستخلاص العناصر الأولية لبنية الكلمة وتحديد سماتها الصرفية.

يوفر استخدام التحليل الصرفي البحث بالمعاني المتعددة التي توجد عليها الكلمة في اللغتين العربية والإنجليزية، ويعتقد الباحث أنه لا يوجد كثير من المستفيدين من يرغب في هذا البحث إذا ما عرف معنى المصطلحات التي يستخدمها لبحث موضوع ما. وقد عثر الباحث عن هذا الشكل في خاصية البحث المتقدم لمحرك بحث عربي هو (عجيب www.ajeb.com)، ومع استخدام كلمة البحث (سيد) خرجت النتائج متضمنة كلمات تحمل معنى السيد أو السيادة بجانب استرجاع الوحدات التي تتضمن كلمة البحث ذاتها.

١٠- التعرف إلى الأخطاء العربية الشائعة:

يقصد بهذا الشكل من البحث التعرف إلى الأخطاء التي توجد عليها كلمات اللغة العربية سواء في بناء النص واسترجاعه. ولغة العربية طبيعة تختلف عن تلك اللغات الغروية أو اللصقية مثل اللغات اللاتينية والإنجليزية. وتعد الهاء الختامية والتاء المربوطة والياء المقصورة وهمزة الوصل والقطع من أكثر الأخطاء شيوعاً بين الكتاب والمستفيدين، ومثال ذلك فإن كلمة (احمد) تسترجع معها الشكل الآخر من الهمزة وهو (أحمد). ويجب في هذه الحالة تعطيل خاصية البحث بحساسية الأحرف؛ لأنها تعمل بشكل عكسي لهذا النوع من البحث.

٢/٢/٤ برمجيات وتطبيقات البحث باللغة العربية:

تحتاج اللغة العربية إلى مجموعة من البرمجيات والتطبيقات التي تسهم في دفع البحث عن مصادر

ثانياً: المصحح الآلي:

النصية الخاضعة لمعالجة حاسوبية وفق منهجية دقيقة ومكثفة وذات مردودية عالية.

تستمد قاعدة المعطيات المعجمية قوتها من المكنز، وهو بذلك يزودها بما يلي:

أ - المعاني الناقصة في المعاجم.

ب - التعابير الاصطلاحية.

ج - كل أنواع الظروف.

د - ألفاظ الحضارة والمصطلحات العلمية الحديثة.

هـ - أنماط مختلفة للإزاحة الدلالية.

و - معلومات عن الحذف العهدي لبعض مكملات الجملة الفعلية.

ز - معلومات عن الصفة السببية.

ح - معلومات قابلية حذف الموصوف الإنسان أو المعنى المجرد.

ط - معلومات صرفية نحوية كالتنوين، أو المنع من الصرف والإفراد والتثنية والجمع والتذكير والتأنيث، وما إلى ذلك.

ي - الأنماط التركيبية للجملة العربية السائدة في الكتابات المختلفة.

ك - أسماء الأعلام العربية والأجنبية.

رابعاً: التشكيل والإعراب الآلي:

لقد أصبح بالإمكان التغلب على أكبر معضلات اللغة العربية المتجلية في الكتابة العربية، بالتمكن

يعتبر «سراج» أحد تلك البرامج العربية التي تسعى صخر لتطويرها في مجال تطبيق منهجيات البحث والتصحيح في النص داخل قواعد بيانات النص العربي. يقوم سراج بعمليات تحليل للنصوص، مثل تصحيح النص واستخلاص كلماته الأساسية وتصنيفه وتلخيصه واستخراج الأعلام^(٢٢).

وهذه الملامح مثل :

- يصحح كلاً من النص العربي والنص الإنجليزي على حد سواء.
- يصحح الأخطاء العربية الشائعة.
- يصحح الأخطاء النحوية العربية.
- يمكنه العمل كمصحح إملائي دقيق دون الحاجة للتدخل البشري.
- يمكنه ربط النص بإحدى الطرق المنهجية للكتابة العربية.
- وضع خاص لتوحيد الأسماء والتصحيح.
- يقوم المصحح الآلي من صخر باكتشاف الأخطاء الإملائية العربية والأخطاء العربية الشائعة وتصحيحها إضافة إلى الأخطاء النحوية.

ثالثاً: قاعدة البيانات المعجمية :

وتتجلى أهميتها في كونها مهياًة ومجهزة لتحليل الأنماط المعجمية، وقدرتها على الإجابة عن كل التساؤلات المتعلقة بالمفردات داخل سياقاتها دلالة وتكريماً، وهو ما يجسد أهمية قاعدة المعطيات

الشريف، ونظام الاسترجاع في قواعد النصوص العربية، والبرامج الكثيرة في مجال التعليم والتكوين والثقافة، بالإضافة إلى كل هذا فإن المشكّل الآلي يوظف في كثير من التطبيقات التقنية التي تهتم بمعالجة النصوص العربية، مثل البحث في الإنترنت، وتحويل النص المكتوب إلى منطوق، وكذلك الترجمة الآلية من العربية وإليها.

خامساً: المصحح النحوي:

ويأتي في السياق نفسه المصحح النحوي، الذي أصبح بإمكانه القيام بتصحيح الأخطاء النحوية التي نجدها شائعة في كثير من الكتابات الصحفية والإنشائية وغيرها، مثل تصحيح إعراب الكلمات التي تلزم حالة الرفع والجر، والأفعال التي تلزم حالة النصب والجزم، وكذلك أخطاء الممنوع من الصرف والمفعول المطلق، وعدم المطابقة النحوية بين الصفة والموصوف والمعطوف والمعطوف عليه، وأخطاء كتابة الأعداد البسيطة والمركبة والمعطوفة وتمييزها، وأخطاء الجملة الفعلية والاسمية، وغيرها من الأخطاء النحوية الأخرى في اللغة العربية.

ويمكن القول في هذا الصدد إن معالجة النحو آلياً على نطاق اللغات الطبيعية قد قطعت أشواطاً في دمج الدراسات اللسانية بالدراسات الحاسوبية وامتزاجهما، وأصبحت العلاقة بينهما علاقة تبادل واقتراض، وأي تنظير للغات الطبيعية أصبح مرهوناً بمعرفة دقيقة للأساليب المنهجية والتحليلية لعلوم الحاسوب، الأمر الذي يؤهل اللغوي لكي يستخدمها

من التشكيل باعتباره من مكوناتها الأساسية وما يلابسها من إشكالات ليس هنا مجال طرحها فيما يمس الكتابة والقراءة والتعليم والوظائف النحوية، وعلى الرغم من كل هذه الإشكالات، ولكون تشكيل الكلمات يؤدي إلى تحديد معانيها وحالاتها الصرفية كان من الضروري إيجاد حلول آلية وتجاوز المشاكل في التعامل مع النصوص العربية في الحواسيب، وهذا بالضبط ما أنجزته صخر بكثير من الدقة تصل إلى (٩٨٪) حيث بإمكان المستخدم لهذا البرنامج أن يحدد نوع التشكيل المطلوب سواء أكان تشكيلاً كاملاً أم تشكيلاً جزئياً لازماً لفك اللبس الصرفي، سواء تعلق الأمر باللغة العربية في نصوصها القديمة أو المعاصرة.

وتعتبر تقنية التشكيل الآلي في ضوء دراسات لسانية إحصائية دقيقة من أهم منجزات صخر في ميدان تطويع اللغة العربية في تعاملها مع الحاسوب، وقد اعتمدت تقنية التشكيل الآلي على عدة مستويات لتحليل ومعالجة اللغة، منها المستوى الصرفي للكلمات، والمستوى الإعرابي، ومستوى التحليل الدلالي، وسمحت هذه التقنية باستخدام أنظمة متطورة جداً في الذكاء الاصطناعي، والاعتماد على المعاجم اللغوية الضخمة، الأمر الذي تطلب سنوات طويلة من البحث الدقيق، ولم يأت برنامج المشكّل الآلي من فراغ، فهو جزء من مشروع ضخم يهدف أساسياً إلى معالجة اللغة العربية آلياً، ولقد ساعد على إنجاز عدد البرامج المتطورة نشير هنا إلى بعضها : برنامج القرآن الكريم والحديث

٤- القيام بفهرسة آلية فعالة لوثائق النص الناتجة عن عملية التمييز.

نتائج الدراسة:

خلصت الدراسة الى مجموعة من النتائج تتمثل فيما يلي :

(١) أن المطلوب تمكين المستخدم العربي من استخدام لغته العربية منذ بدء تشغيل جهاز الحاسب حتى الوصول إلى أي معلومة على الإنترنت. ولكن توجد حالياً مشكلة تحد من استخدام اللغة العربية. تنحصر في أن الوضع الحالي يفرض على المستخدم العربي حينما يود الوصول إلى المواقع، وخاصة العربية منها أن يدخل عنوان الموقع بالأحرف اللاتينية حتى وإن كان المحتوى باللغة العربية، وهذا يعد من المعوقات الرئيسة لانتشار الإنترنت في العالم العربي، لذلك كان من البديهي الحاجة إلى تعريب أسماء المواقع، بحيث يتمكن المستخدم العربي من استخدام اللغة العربية للوصول إلى المعلومة ببسر وسهولة.

(٢) عمل محرك البحث Google على دعم اللغة العربية وقدرات البحث بالمصطلحات العربية، غير أن هذا الدعم لم يكن على مستوى الملاءمة التامة لخصائص وطبيعة اللغة العربية التي تتعدد على مستوى الكلمة الواحدة أو العبارات المركبة. ومن ثم فإن محرك البحث بحاجة الى ضم برمجيات وتقنيات متطورة تلائم ملامح

في صياغة جديدة للمنظومة النحوية، وتطوير علم الدلالة وكل ما يتعلق أيضاً بالصناعة المعجمية.

سادساً: القارئ الآلي O C R :

يعد القارئ الآلي إحدى التقنيات المتطورة جداً في مجال الصناعة الإلكترونية المعلوماتية وجعل اللغة العربية تحتل مكانتها اللائقة بها بين اللغات الطبيعية، إذ أصبح بالإمكان التعرف إلى النصوص العربية بمختلف أشكال كتابة حروفها، سواء كانت مشكلة أو غير مشكلة، حيث تحول إلى ملفات إلكترونية قابلة للمعالجة والتصرف فيها، بالإضافة أو الحذف، ولقد أصبح القارئ الآلي يملك قدرة التعرف إلى حروف الكتابة العربية بجميع أشكالها بسرعة فائقة تزيد على ٥٠٠ حرف في الثانية الواحدة، وبالإضافة إلى ذلك يتيح إمكانية تخزين النصوص وحفظها بشتى أنواع الصيغ المعروفة، وحتى صيغة HTML و UNICOM.

ومن بين خصائص القارئ الآلي:

- ١- قدرته على تمييز عدد كبير من اللغات وتعريفها بواسطة أدوات التمييز المتاحة، فبالإضافة إلى اللغة العربية والفارسية فإنه يدعم إحدى عشرة لغة أوروبية، وهي: الإنجليزية والفرنسية والألمانية والإيطالية و الإسبانية والبرتغالية والنرويجية والسويدية والدنماركية والنمساوية والفرنلندية.
- ٢- احتواؤه على ٢٦ مكتبة خطوط عربية.
- ٣- تحديد الصور في الملف آلياً وإخراجها مع النص الناتج.

محمود حيث يجبر المستخدم على كتابة العنوان بلغتين مختلفتين في اتجاهين متعاكسين. لذا لا بد من أن تكون جميع أجزاء اسم النطاق باللغة العربية بما في ذلك النطاقات العليا.

(٦) أن استخدام أسماء الإنترنت باللغة العربية هو من باب تشجيع المستخدم العربي على استخدام الإنترنت وكسر حاجز اللغة، حيث إن استخدام اللغة العربية لأسماء المواقع يسهل الوصول إلى المواقع. فإن تعريب المحتوى يعتبر خطوة جيدة، ولكن حتى تكتمل الفائدة فلا بد أيضاً من استخدام العنوان باللغة العربية حتى يتمكن المستخدم العربي من الوصول إلى المحتوى العربي. وحيث إن التوجهات الحالية منصبّة حول الحكومة الإلكترونية والتجارة الإلكترونية، لذلك فالجهات، سواءً كانت حكومية أو تجارية، والتي تود تطبيق الأعمال الإلكترونية في البلدان العربية لا بد لها من مخاطبة المواطنين والمستهلكين بلغتهم وأن تجعل مواقعها على الإنترنت في متناول من لا يحسن اللغة الإنجليزية.

(٧) أن الجهات المطورة والمنتجة لتقنيات تعريب استخدام الأسماء على الإنترنت مطالبة بالاتفاق فيما بينها على معايير ومقاييس أساسية وتمريها من خلال القنوات المسؤولة عن إصدار المعايير القياسية للإنترنت. ويجدر التنبيه والتأكيد على أهمية التنسيق بين الجهود العربية والجهات الأخرى لوضع ضوابط

الكلمات العربية، مثل المحلل الصرفي والمصحح النحوي، وذلك لتلافي الوقوع في مشكلات اللغة مثل الاشتراك اللفظي والترادف والبحث بالمشتقات... وغيرها.

(٣) أن من يتحدثون اللغة الإنجليزية في الوطن العربي لا يتجاوز تقريباً ١٠٪. وهي نسبة ضئيلة جداً إذا ما قارناها بالعدد الكبير من المستخدمين المتوقع دخولهم إلى عالم الإنترنت، لذا تبقى اللغة حاجزاً يمنع انتشار استخدام الإنترنت بين العرب ما لم توفر المادة والسبيل (العنوان) بلغة المستخدم. وأيضاً ستكون اللغة حاجزاً أمام الاستفادة من الإنترنت في التعليم للمراحل الأولى، حيث إن معظم الطلبة في العالم العربي يبدؤون تعلم اللغة الإنجليزية في مراحل التعليم العليا.

(٤) تعد اللغة العربية من أعرق اللغات وأسمائها تاريخاً، وأطولها باعاً في العلوم؛ لذلك يتحتم على الأمة العربية والإسلامية المحافظة عليها وعدم استبدالها بأي لغة كانت تحت أي ظرف، والعمل على تطوير التقنية والحلول الفنية لدعم لغتنا لغة القرآن الكريم، وأن نخدمها كما خدمها أسلافنا الأولون.

(٥) هناك توجه نحو استخدام اللغة العربية لاسم الجهة وإلحاقه بإحدى النطاقات العليا (TLDs) الممثلة باللغة الإنجليزية (مثل: اسم- شركة.com)، وهذا توجه غير منطقي وغير

(د) الحلول الفنية لدعم استخدام اللغة العربية على الإنترنت.

(٨) أثبتت الدراسة أن تخصيص الموازنات الكافية والكبيرة وتفرغ العلماء الأجلاء لخدمة هذه اللُّغة التي تراكمت وزاد اليابس في أعضائها فاحتاجت إلى تقليم وتشذيب. وربط اللُّغة بالمعارف العلميّة وتأثير هذه المعارف عليها، وجعلها تأخذ مكانتها في الاستعمال اللُّغويّ المرن، وليس تطويع المعارف للغة، كما عملت بعض المجامع العربيّة، أو صناعة مفردات لم تستعمل؛ فالمعارف تاريخياً كانت تفرض مفرداتها ومضامينها الجديدة، والذين يغلّقون أبواب اللُّغة عن الاستعمال والتّطوّر من أبنائها إنّما يحاولون خنقها ومنع الهواء والتّنفّس عنها بحجّة حمايتها.

ومعايير ثابتة تتوافق مع المعايير الدولية، حيث إن الرغبة باستخدام لغات محلية غير الإنجليزية هي عامة وتهم جميع أصحاب لغات العالم الحية. ودعم اللغة العربية في كتابة أسماء المواقع يحتم دراسة الموضوع من عدة جوانب، منها:

(أ) وضع مقاييس موحدة لتعريف مجموعة المحارف العربية المسموح باستخدامها في كتابة أسماء النطاقات العربية.

(ب) وضع مقاييس موحدة لهيئة هيكل الأسماء العربية (شجرة أسماء الإنترنت العربية)، بما في ذلك تحديد النطاقات العربية العلوية العامة (general TLDs gTLDs) والدولية (country code TLDs – ccTLDs).

(ج) تنظيم خادمت أسماء النطاقات الرئيسة (DNS root servers) الخاص باللغة العربية.

الهوامش والمراجع

- (٣) محركات البحث العربية : إنترنت العالم العربي . - ع مايو / أغسطس (١٩٩٨م) . تاريخ الاطلاع مارس ٢٠٠٢م . - متاح في : www.ditnet.co.ae/arabic/internet/studyd0110_3.html
- (٤) محركات البحث العربية : إنترنت العالم العربي، المصدر السابق نفسه.
- (٥) أريزونا عبد العزيز أبنامي . خطوات بسيطة تضمن الاستفادة القصوى من محركات

- (١) عادل محمد أحمد خليفة. التحول إلى النشر الإلكتروني : حلول واقعية، تاريخ الاطلاع ١٠/١/٢٠١١م، متاح على www.arab-pa.org/.../النشر-٢٠٪ الإلكتروني. ٢٠٪-٢٠٪-٢٠٪عادل٢٠٪خليفة. doc
- (٢) المحتوى العربي على شبكة الإنترنت. تاريخ الاطلاع ١٢/١/٢٠١١م، متاح على <http://www.arabiancreativity.com/fadl4.htm>

في اللغة العربية -. الرياض : مكتبة الملك فهد

الوطنية، ١٩٩٤م . ص ٦٨-٧٧

(١٤) محمد سالم غنيم. الحالة الراهنة لنظم

استرجاع المعلومات العربية المعتمدة على اللغة

الطبيعية دراسة في أدب الموضوع، مجلة مكتبة

الملك فهد الوطنية، مج ١٣، ٢٤، رجب - ذو

الحجة ١٤٢٨هـ / يوليو.

(١٥) سيد ربيع سيد إبراهيم. محركات بحث الصور

الثابتة على الإنترنت : دراسة تحليلية. الرياض:

مكتبة الملك فهد الوطنية، ٢٠٠٧م.

(١٦) إرشادات البحث في الإدرسي. ١٠/١٠/٢٠٠٢م،

مسار <http://asb.kacst.edu.sa/>

IDRADMIN/Arabic/Help/ATips.htm

(١٧) مساعد بن صالح الطيار. محركات البحث

العربية وطرائق البحث فيها، الجزيرة،

٢٠٠١م، تاريخ الاستشهاد ١٠/١٠/٢٠٠٢م،

مسار [http://www.al-jazirah.com.sa/](http://www.al-jazirah.com.sa/evillage/22112001/wr1.htm)

evillage/22112001/wr1.htm

١٨ إرشادات البحث في الإدرسي. مصدر سابق، ص ٤.

(١٩) www.alidrissi.com

(٢٠) عبد الغني أبو العزم. اللغة العربية والمعالجة الآلية

برامج صخرنموذجاً، تاريخ الاطلاع ١٢/٤/٢٠١١م،

متاح على <http://www.aljabriabed.net/>

n31_04abualazm.(2).htm

(٢١) عبد الغني أبو العزم. المصدر السابق نفسه.

(٢٢) المصحح الآلي "سراج". تاريخ الاطلاع

www.ccse. ١/٥/٢٠١١م، متاح على

[kfupm.edu.sa/~husni/Teaching/](http://kfupm.edu.sa/~husni/Teaching/ACStLect13.doc)

ACStLect13.doc

البحث. - جريدة الجزيرة . - ع ٩٩٩٣ (فبراير

٢٠٠٠م). - تاريخ الاطلاع مايو ٢٠٠٢م. - متاح

في www.suhuf.net.sa:

(٦) أمنية طلعت . الطريق الآمن لدخول عالم

الإنترنت . - جريدة البيان . - ٢٧ يناير

٢٠٠٢م. - تاريخ البحث مارس ٢٠٠٢م. -

متاح في : www.albayan.co.ae/

albayan/2002/01/271mnw/9.html

(٧) ماهية البحث على الإنترنت. تاريخ الاطلاع

<http://ccisdb>. متاح على

ksu.edu.sa/files/rep9900000.doc

(٨) ماهية البحث على الإنترنت. المصدر السابق نفسه.

(٩) دليلك إلى محركات البحث في الشبكة : إنترنت

العالم العربي .- ع مايو / أغسطس (١٩٩٨م).

تاريخ الاطلاع مارس ٢٠٠٢م. - متاح في

[www.ditnet.co.ae/arabic/internet/](http://www.ditnet.co.ae/arabic/internet/studyd0110_2.html)

studyd0110_2.html

(١٠) نبيل بدر سنبل . مقدمة إلى محركات البحث

في الشبكة . - الرياض : جامعة الملك فهد ،

٢٠٠١م. - تاريخ البحث مايو ٢٠٠٢م. - متاح

في <http://Users.kfupm.edu.sa>:

(١١) عبد الرحمن بن حسن المحسني. قضية

الترادف : النظرية والتطبيق، تاريخ الاستشهاد

<http://www>. مسار ٢٢/١٠/٢٠٠٢م،

adabiabha.com/biader/4.htm

(١٢) المصدر السابق نفسه.

(١٣) علي السليمان الصوينع. استرجاع المعلومات