

أدوات وتقنيات البحث في الويب/د.كداوه عبدالقادر

أدوات البحث في الويب:

تتوفر في الويب مجموعة من الأدوات التي تمكن الباحث من الإبحار في محتوياتها وتشكل هذه الأدوات واسطة بين الباحث ومحتوى الويب ومن أهمها نجد محركات البحث والأدلة الموضوعية ومحركات البحث الفائقة.

➤ محركات البحث:

محركات البحث هي أدوات بحث تعمل من خلال استراتيجيات محددة (البحث البولوني) أو إستراتيجيات مفتوحة (البحث باللغة الطبيعية) وذلك للبحث في حقول أو وثائق نصية والأكثر من ذلك أنها يمكن أن تبحث عن أشياء (صور ورسوم وخرائط وأصوات)، في بيئة محددة هي بيئة شبكة الانترنت، وذلك يعني أنها تبحث في ملايين المواقع ومليارات الكلمات في وقت محدد وتتميز بسرعة الاستجابة وعادة ما تكون إجابتها مواقع على الانترنت تتوافر فيها كل المصطلحات التي تم البحث عنها أو بعضها، أو مواقع محددة سلفا من خلال ما يعرف بقائمة أو دليل البحث.

آلية عمل محرك البحث:

محركات البحث المعروفة حاليا رغم وجود اختلاف بينها كما هو الحال في طريقة ترتيب نتائج البحث الا انها تعتمد على نفس الية العمل ويمكن تقسيم هذه الآلية إلى أربع خطوات أساسية وهي:

الخطوة الأولى: جمع المعلومات

هذه المرحلة يتم من خلالها تجميع وتكشيف المعلومات المتواجدة في صفحات الويب المختلفة من خلال ما يعرف بربوات التكشيف- العناكب والديدان-

الخطوة الثانية: التكشيف وبناء الكشاف

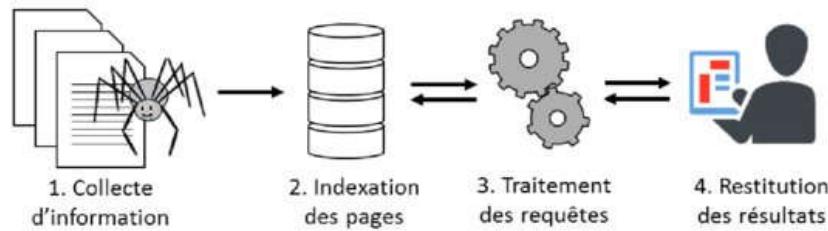
هذه المرحلة من خلالها يتم تخزين المعلومات ناتج عملية التكشيف- المرحلة السابقة- في قاعدة بيانات خاصة بكل محرك بحث- الكشاف-

الخطوة الثالثة: معالجة معادلات البحث

في هذه المرحلة تتم معالجة معادلات البحث المقترحة من طرف الباحث وفق الطرق والمعايير المستخدمة من قبل محرك البحث والمضاهات بينها وبين محتوى قاعدة البيانات الخاصة بمحرك البحث.

الخطوة الرابعة: استرجاع النتائج

في هذه المرحلة الأخيرة يتم إظهار النتائج المتوصل إليها في المرحلة السابقة ويتم عرضها في الواجهة الخاصة بمحرك البحث وتحتوي على مجموعة من الروابط التي تحيل الباحث إلى الصفحات التي تحتوي على المعلومات المتعلقة بمعادلة البحث المقترحة من طرفه. والشكل الموالي يوضح آلية عمل محرك البحث:



شكل 1: آلية عمل محرك البحث

من خصائص محركات البحث أنها تمكن من الوصول إلى كم هائل من النتائج والتي تتميز بالحدثة لأنها تعمل بشكل آلي ولكن قد تكون بعيدة عن موضوع البحث لأنها تعتمد على المستوى الشكلي للغة ولا يتدخل في عملها العقل البشري.

➤ الأدلة الموضوعية:

تعريف الدليل الموضوعي: هي مواقع متخصصة على الانترنت تنتقي مواقع ويب أخرى وتنظمها تحت رؤوس موضوعات واسعة مثل الفن والترفيه والتسلية والعلوم. ويمكنك أن تتصفح باعتماد موضوعات عريضة إلى أن تجد الموضوع المحدد الذي ترغب فيه ، أو من خلال البحث ضمن الدليل الموضوعي باستخدام الكلمات المفتاحية.

الأدلة الموضوعية تعتمد في عملها على العقل البشري بينما محركات البحث تعمل بشكل آلي بحث.

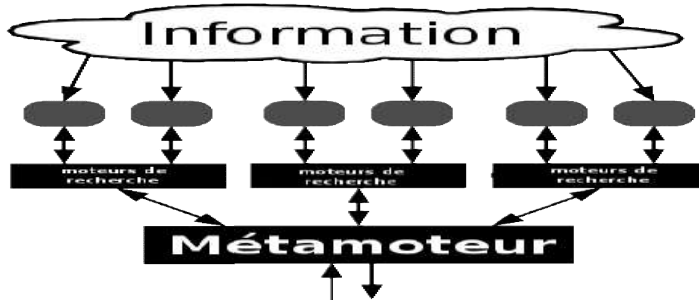
من بين ايجابيات الأدلة الموضوعية مايلي:

- ✓ يعتمد البحث على السياق: كما في ترتيب المكتبة بواسطة الموضوع فان المواد ترتب مع المواد المماثلة والتصفح الجانبي قد يعرض إمكانيات مثمرة.
 - ✓ مصادر مختارة: الانتقاء قد يعوض نقص الشمولية في بعض الأحوال.
 - ✓ تحسين فرص إيجاد مصادر ذات جودة عالية: بسبب مراقبة الإنسان في صيانتها الأدلة الموضوعية فهي تملك قدرة توصيل عالية الجودة.
 - ✓ مخاطر اقل في التكرار والوفرة: عكس الكشافات الآلية فان الأدلة الموضوعية لا توفر عادة العديد من الإصابات لنفس العمل.
 - ✓ كما أنها أيضا لا توفر العديد من النتائج الخارجة عن سياق الاستفسار مثل محرك البحث.
- و من عيوب الأدلة الموضوعية نجد:

- ✓ البحث في التسلسل الهرمي قد يكون مستهلكا للوقت.
 - ✓ يجب أن يكون هناك فكرة عن كيفية تصنيف الصفحات.
 - ✓ تشير الأدلة إلى الصفحات فقط ومحتوى الصفحات يمكن أن يتغير دون ان يدرك المحررون ذلك.
 - ✓ نظرا لأنها تعتمد على مجهود المكشفين فهي لا تملك قاعدة بيانات ضخمة وتغطيتها ضعيفة مقارنة بمحركات البحث.
- محركات البحث المتعددة:

محرك البحث المتعدد هو محرك بحث يبني إجاباته على محركات أخرى ويحلل النتائج قبل تسليمها لمستخدمي الإنترنت. محرك البحث الفائق (أو محرك البحث المتعدد) هو محرك بحث يستمد معلوماته من عدة محركات بحث عامة. بتعبير أدق ، يرسل محرك البحث الفائق طلباته إلى العديد من محركات البحث ويعيد نتائج كل منها. يسمح محرك البحث الفائق للمستخدمين بإدخال موضوع البحث الخاص بهم مرة واحدة فقط أثناء الوصول إلى الردود من عدة محركات بحث مختلفة.

ومن بين محركات البحث الفائقة نجد: Yasni YouCare WebCrawler Startpage Searx Lilo Ecosia DuckDuckGo والشكل الموالي يوضح بنية محرك البحث الفائق والية عمله.



الشكل 2: آلية عمل محرك البحث المتعدد

مزايا محركات البحث المتعددة:

من مزايا محركات البحث المتعددة:

- ✓ تحفظ وقت البحث، كما أنها طريقة سريعة لتحديد المحركات التي تسترجع أفضل النتائج لبحث معين.
- ✓ تجنب المستخدم تعلم أساليب البحث المختلفة لكل أدوات البحث المتاحة.
- ✓ تكنولوجيا متطورة لمعالجة مشكلة عدم قدرة أداة البحث الواحدة من تغطية الويب.

تقنيات البحث في شبكة الانترنت:

• المنطق البولياني:

المنطق البولياني أو الروبوت المنطقية او- و- الا AND-OR-NOT تعتبر من أهم تقنيات البحث في الويب والتي تساعد الباحث في ضبط معادلة البحث. (Lan, 1999 ,p104)

- تستخدم AND -و- لربط المصطلحات التي نبحث عنها والقيام ببحث أكثر دقة بحيث نتحصل على كل المصطلحات المعنية بالبحث في نفس الوقت، على سبيل المثال: التلوث AND البحر AND السماء. كل نتائج البحث في هذه الحالة تحتوي على الكلمات الثلاث في نفس الوقت.
- نستخدم OR-او- لتوسيع نطاق البحث والحصول على إحدى الكلمات التي نبحث عنها، وتستخدم للبحث عن المرادفات، على سبيل المثال: التلوث OR البحر OR السماء. نتحصل في هذه الحالة على عدد كبير من النتائج والتي تحتوي على الأقل كلمة من هذه الكلمات الثلاث.
- نستخدم NOT-إلا- لإقصاء كلمات من البحث، على سبيل المثال التلوث NOT البحر، الي النتائج تكون كل ما يتعلق بالتلوث ما عدا البحر.
- يمكن استخدام عدة عوامل منطقية مع بعض من خلال الاستعانة بالأقواس مثلا: التلوث AND (البحر OR الأرض) للحصول على نتائج حول التلوث في البحر او التلوث في الأرض.
- يمكن ان تعوض برموز رياضية مثلا + تعني AND و - تعني NOT و / تعني OR
- علامات التنصيص "" :

تستخدم للبحث عن عبارة بالضبط مثلا "البحث التوثيقي الآلي" في هذه الحالة كل نتائج البحث الظاهرة تحتوي على نفس العبارة.

• البحث عن نوع معين من الملفات:

بعد إضافة الامتداد الخاص بكل نوع من الملفات للكلمة أو العبارة التي نبحث عنها النتيجة تكون ملفات من هذا النوع مثلا: pdf,doc ...

• تقنية البتر:

يمكن للباحث استخدام * للتعبير عن مجموعة من الحروف أو العبارات غير معروفة تقع بين كلمتين أو تكون سابقة أو لاحقة لجذر كلمة على سبيل المثال: جيو* للحصول على جيولوجيا، جيوفيزياء،.... أو مثلا: * لوجيا للحصول على: بيولوجيا، جيولوجيا، انطولوجيا،.... الخ.

• تقنيات خاصة بمحرك البحث جوجل:

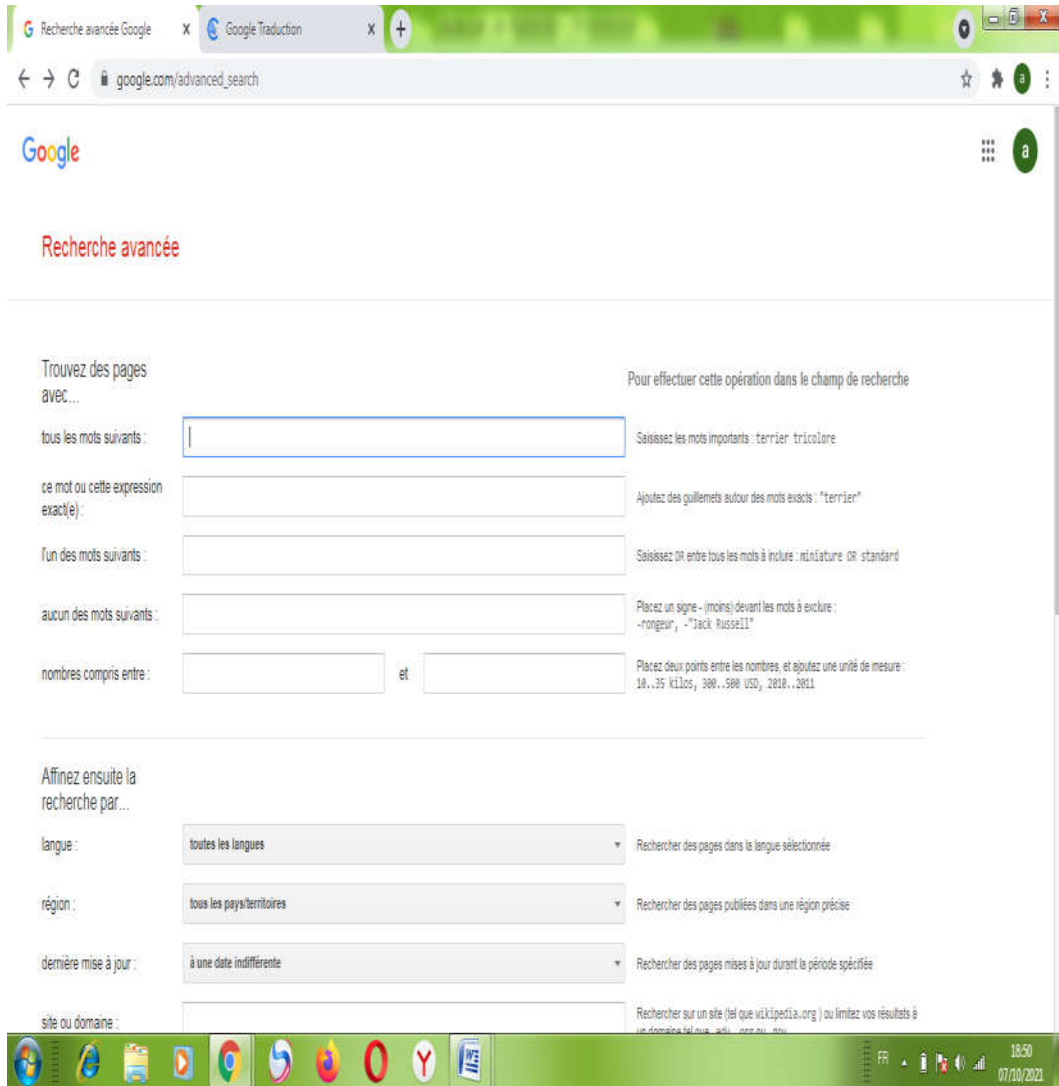
جدول 2: تقنيات البحث الخاصة بمحرك البحث "جوجل"

	للبحث في:
Intitle :	عنوان الصفحة:
Inurl :	عنوان الموقع:
Site :	الموقع:
Site :	الميدان (النطاق):
Define :	لتعريف مصطلح معين:

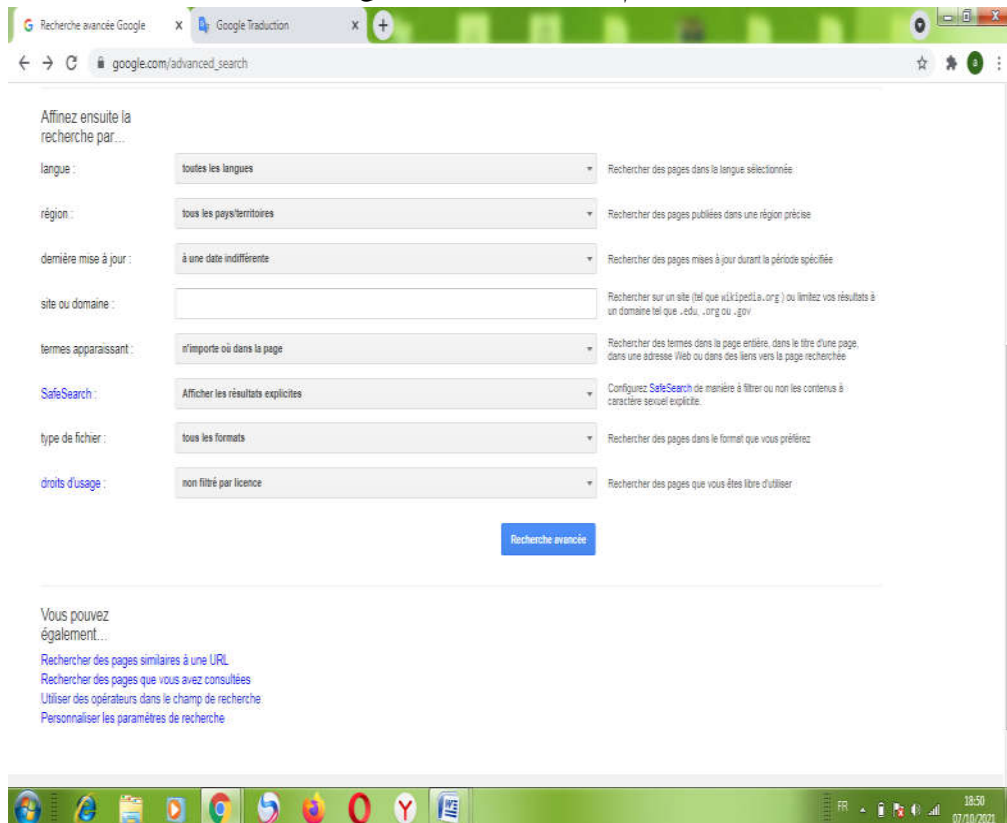
ومعظم هذه التقنيات نجدها ملخصة في خاصية البحث المتقدم كما هو الحال بالنسبة لمحرك البحث جوجل كما هو موضح في الشكلين الموالين:

البحث المتقدم في جوجل يتيح للمستفيد العديد من الخيارات كما هو موضح في جزئين:

- الجزء الأول: مجموعة من الخيارات التي تختصر دور العوامل المنطقية وعلامات التنصيص مثلا الإطار الأول كل هذه الكلمات (AND) والإطار الثاني البحث عن عبارة بالضبط ("") والإطار الثالث البحث عن إحدى الكلمات (OR) والإطار الرابع خاص بالاستثناء (NOT)
- الجزء الثاني: يتيح للباحث اختيارات أخرى لضبط معادلة البحث مثلا اختيار اللغة والمنطقة واخر تحديث للموقع ونوعه ونوع الملف pdf,doc,ppt... وغيرها من الاختيارات.



الشكل 3 : واجهة البحث المتقدم لمحرك البحث جوجل ج1(www.google.com)



الشكل 4 : واجهة البحث المتقدم لمحرك البحث جوجل ج2(www.google.com)