الانترنت والاتصالات

Internet & communications

LAB #6

أستاذة المقرر :
البريد الإلكتروني :

محركات البحث (search engines):

- هي أدوات شائعة الاستخدام للوصول إلى صفحات النسيج الشبكي العالمي World Wide Web (WWW).

- لكنها تعود بمئات النتائج والوصلات التي تكون غالباً ليس لها صلة بموضوع البحث الرئيسي مما يجعل الباحث في حيرة من ما يبحث عنه. ويعود السبب في ذلك إلى استخدام أدوات البحث المتوفرة على الانترنت غالباً بشكل عشوائي وبدون إستراتيجية بحث محددة.

أدوات البحث :

- يستخدم مستخدمو الانترنت عدة أدوات للبحث على الويب للوصول إلى المعلومات المفيدة الموزعة على ملايين الأجهزة الخادمة على الانترنت والتي تساعد في تسريع عملية البحث بدرجة كبيرة .

ما الفرق بين
محركات البحث والأدلة؟

محركات البحث (Search Engines) :

تعريفها :

عبارة عن قواعد بيانات ضخمة بعناوين ومواقع ، ومع وصف مصغر لصفحات الانترنت المختلفة ، والتي بواسطتها " أي محركات البحث " يمكن البحث عن موضوع معين في حقل من الحقول المختلفة في الشبكة بشكل دائم بغرض إيجاد دليل لمثل هذه الصفحات .

مميزاتها:

١. تعمل بشكل آلي .
٢. تقوم بفرز وفهرسة كم هائل من الصفحات.
وبسبب هاتين الميزتين ، ظهرت لها ميزة ثالثة ..
٣. تحتوي على كثير من المعلومات غير المتوفرة في الأدلة مما يجعلها أداة فعالة أكثر من الأدلة .

- من أشهر محركات البحث : محرك بحث Google

الأدلـــــة (Directories) :

طريقة عملها :

- لا تعمل بشكل آلي بل تتم إدارتها من قبل أشخاص متخصصين ، وما يحدث هو أن العديد من المواقع يتم تسليمها إلى دليل ما ، ثم يتم فرزها وتبويبها تحت تصنيف معين .
- تقدم الأدلة للمستخدم طريقة سريعة للبدء بعمليات البحث بواسطة تفحص المواضيع المصنفة التي يعرضها :
  - إذ يندرج تحت كل موضوع لائحة من المواضيع الفرعية فيمكن للمستخدم أن يتفحصها تباعاً إلى أن يصل إلى المعلومات المطلوبة .
  - في حال عدم وجود المعلومات تحت الموضوع الذي اختاره المستخدم ، يتراجع ويختار موضوعاً رئيسياً آخر ليقوم بالبحث في تفرعاته من جديد وهكذا ..

الأدلـــــة (Directories) :

مميزاتها:

١. تقدم للمستخدم طريقة سريعة للبدء بعمليات البحث بواسطة تفحص المواضيع المصنفة التي يعرضها.
٢. يتم إدارتها بشكل بشري .
لذلك ..
٣. هي قادرة دوماً على توفير معلومات أكثر دقة .

- من أشهر الأمثلة على الأدلة : موقع yahoo
- هناك العديد من أدلة البحث تعمل أيضاً كمحركات منها :
  - Yahoo
  - AltaVista
  - excite

مكونات محركات البحث وكيفية عملها:

- يوجد على شبكة الانترنت كميات ضخمة من المعلومات تتوزع على شكل صفحات وتنمو بشكل سريع .
- يقدر عدد الصفحات التي تضاف أو تحدث يومياً بمليون ومائتين صفحة.
- وقد أثبتت الدراسات أن معدل نمو الانترنت يتضاعف بشكل سنوي .
- هناك أكثر من ٢٥٠٠ أداة بحث متوفرة .
- تقدر تغطية تلك الأدوات بـ:

٥% للأدوات الصغيرة	١٥% للأدوات العملاقة

- ولذلك فإنه مهما كانت سمعة أداة البحث فإنها لا يمكن أن تغطي كافة المعلومات المتوفرة على الشبكة.

مكونات محركات البحث وكيفية عملها:

- تُظهر محركات البحث نتائج مختلفة ! (لماذا ؟) ..
- لأن كل منها يبحث في قاعدة بياناته الخاصة التي تتضمن جزء من مواقع الويب .

- يتألف محرك البحث من ثلاثة أجزاء رئيسية وهي :
  ١. برنامج العنكبوت Spider Program .
  ٢. برنامج المفهرس Indexer Program .
  ٣. برنامج محرك البحث Search Engine Program .

١- برنامج العنكبوت (Spider Program):

- تستخدم محركات البحث برنامج العنكبوت لإيجاد صفحات جديدة على الويب وإضافتها ، ويسمى هذا البرنامج أيضاً بالزائر لأنه يبحر في الانترنت بهدوء لزيارة صفحات الويب والإطلاع على محتوياتها .

- يأخذ هذا البرنامج مؤشرات المواقع من :
- عنوان الصفحة Page Title.
- الكلمات المفتاحية Keywords التي تحتويها الصفحة.
- محتويات محددات الميتا Meta Tags في الصفحة.

١- برنامج العنكبوت (Spider Program):

- **الكلمات المفتاحية Keywords**

هي الكلمة أو الكلمات التي يدخلها الزوار في مربع البحث في محركات البحث وعندما يبدأ محرك البحث بجلب نتائج بحث للكلمة المفتاحية فانه يبحث في محتوى المواقع عن هذه الكلمات وعند ظهور هذه الكلمات في محتوى الموقع أو في عنوان الصفحة أو في أي مكان في الموقع فانه يقوم بإدراج الموقع في نتائج البحث .

- **محتويات محددات الميتا Meta Tags**

- تستخدم واصفات الميتا الخاصة من أجل وصف الموقع و الكلمات المفتاحية في صفحة HTML حتى يتم إدراج المواقع بشكل سليم في محركات البحث. وهي لا تظهر بشكل مرئي على صفحتك ولكنها تعمل في الكواليس حيث تقدم وصفاً عن المعلومات التي تتضمنها الصفحة للمتصفحات وعناكب الشبكة.
- وعندما يزور عنكبوت الشبكة صفحة لا تحتوي على وصف لها باستخدام محدد الميتا فإنه يسجل الأسطر القليلة الأولى من النص الموجود في بداية الصفحة و يعتبر ذلك وصفاً لها.

هذا الأسلوب لا يعطي وصفاً دقيقاً للصفحة.. لماذا؟؟

- السبب هو وجود صورة في بداية بعض الصفحات و برنامج العنكبوت لا يتعرف إلا على النصوص.

١- برنامج العنكبوت (Spider Program):

- لا تقتصر زيارة برنامج العنكبوت على صفحات الموقع ، بل يتابع تعقب الروابط (links) الموجودة فيها لزيارة صفحات أخرى .
- **ما الغاية من هذه الزيارات ؟**
- الغاية هي وضع النصوص المتاحة على تلك المواقع على فهارس محرك البحث ليتمكن المحرك من العودة إليها فيما بعد .
- **ملاحظة** : يتم تغيير محتوى المواقع بين فترة وأخرى ، فهل سوف يحتفظ محرك البحث بالنسخة القديمة من محتويات المواقع للأبد **؟!**
- لا ، بل يقوم محرك البحث بزيارات دورية للمواقع الموجودة في الفهارس للتأكد من التعديلات التي تصيب المواقع المفهرسة.

٢-برنامج المفهرس Indexer Program :

- يطلق عليه أحيانا الكتالوج catalogue.
- هو قاعدة بيانات ضخمة تقوم بتوصيف صفحات الويب .
- يركز التوصيف على المعلومات التي حصل عليها من برنامج العنكبوت .
- تعتمد على بعض المعايير مثل الكلمات الأكثر تكراراً.
- تختلف محركات البحث عن بعضها في :

١. المعايير التي تعتمد عليها في الحصول على المعلومات .
٢. خوارزميات البحث searching algorithms .

٣- برنامج محرك البحث Search Engine Program :

- يبدأ دوره عند كتابه كلمة مفتاحية في مربع البحث search box.
- يأخذ الكلمة المفتاحيه ويبحث عن صفحات الويب التي تحقق الاستعلام الذي كونه برنامج المفهرس في قاعدة بيانات الفهرس.
- ثم يعرض نتيجة البحث ممثلة بصفحات الويب في نافذة المستعرض browser window ويقوم بعملية الترتيب لهذه الصفحات.
- تختلف محركات البحث عن بعضها في أسلوب العمل:

فمثلا تحتفظ قاعدة بيانات AltaVista بكل تفاصيل صفحة الويب المخزنة عليها ،

أما محركات البحث الأخرى فقد تحتفظ بالعناوين الرئيسية للصفحة فقط ، مما يؤدي إلى اختلاف شكل ودقه نتائج البحث الظاهر للمستخدم.

نسبة المواقع التي تغطيها محركات البحث من مجموعة مواقع الشبكة

- أكدت دراسة أجريت على ١١ محرك بحث مستخدمة في الانترنت ، إن هذه المحركات لم تتمكن من التأقلم مع النمو الكبير لمواقع الويب.
- كما بينت الدراسة أن محركات البحث Search Engines لا تغطي سوى قسم ضئيل من صفحات الويب ، وأن إدراج أي صفحه جديدة بينها يتطلب وقتاً طويلاً .
- وجدت الدراسة إن محرك البحث northern light هو أكثر المحركات شمولاً على الرغم من أنه لا يغطي سوى ١/٦ من صفحات الانترنت في وقت الدراسة. وتطورت هذه النسبة بعد مدة من الدراسة لتصل تغطيته إلى ١/٣ من صفحات الانترنت.
- من المحتمل أن يكون تأخر إدراج الصفحات الجديدة ضمن محركات البحث أمراً مقصوداً، وذلك لأنه عندما تمتلئ قواعد البيانات التابعة لمحركات البحث هذه، تصبح عملية إضافة موقع جديد إليها عملية مكلفة، خصوصاً أن هذه الإضافة لا تزيد من الأرباح .

نسبة المواقع التي تغطيها محركات البحث من مجموعة مواقع الشبكة

أسم محرك البحث	نسبة التغطية	دعم اللغة العربية	عنوان محرك البحث
NORTHERN LIGHT	%١٦,٠	لا	WWW.NORTHERNLIGHT.COM
SNAP	%١٥,٥	لا	WWW.SNAP.COM
ALTAVISTA	%١٥,٥	نعم	WWW.ALTAVISTA.COM
HOT BOT	%١١,٣	لا	WWW.HOTBOT.COM
MICROSOFT	%٨,٥	لا	WWW.MSN.COM
INFOSEEK	%٨,٥	نعم	WWW.INFOSEEK.COM
GOOGLE	%٧,٨	نعم	WWW.GOOGLE.COM
YAHOO	%٧,٤	لا	WWW.YAHOO.COM
LYCOS	%٢,٥	لا	WWW.LYCOS.COM
EUROSEEK	%٢,٢	نعم	WWW.EUROSEEK.COM
EXICIT	%٥,٦	نعم	WWW.EXCIT.COM
ALL THE WEB	%١٥,٥	نعم	WWW.ALLTHEWEB.COM

آليات البحث في محركات البحث

١- البحث البسيط

Simple Search

وهو النوع الذي يقوم به معظم الناس حيث يتم وضع كلمة البحث بدون أي علامات أو إشارات.

٢- البحث المتقدم

Advanced Search

تعتبر طريقة فعالة للبحث عن معلومات محددة إذ تتيح للمستخدمين إمكانية البحث عن عدة كلمات مفتاحية معاً باستخدام بعض المعاملات مثلAnd, Or, Not, Near