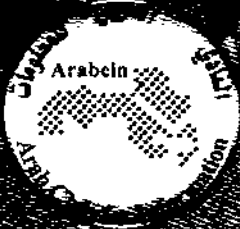


العدد 4-2001



أمين عام
عبد القادر البشير
واسم جرجيس
رمضان أبو خضوة
علي أبو الزهاك
علي صالح كزار
علي طوقان
علاهم رايم
عيسى العسافين
قيليب نبهان
الكبير بطيوة
لطفي كرموس
محمد جميد الجابري
محمد فتحي عبد الهادي
محمود أحمد إتييم
مختار بن هندة

العربية 3000

تصليحة - محكمة - متخصصة

تبحث في قضايا وأفاق علم المعلومات في الوطن العربي
تصدر عن النادي العربي للمعلومات

الرئيس المدير المسؤول

د . عبد المجيد الرفاعي

رئيس التحرير

د . حسنة محيي الدين

نائب رئيس التحرير

نبيل مخول

مدير التحرير

نزار حميد

سكرتير التحرير

هناء غانم

توجه المراسلات باسم هيئة التحرير

إلى العنوان التالي: سورية - دمشق - النادي العربي للمعلومات

ص.ب. 33092 فاكس: 2127648 هاتف: 2128738
5914521 5914521

البريد الإلكتروني: arabcin@net.sy

البحث الذكي وخدمات القيمة المضافة للمعلومات على شبكة الإنترنت

د. المختار بن هنده*

المستخلص :

أمام التدفق السريع للمعلومات في عصر التقنيات الحديثة تغير الإطار العام لمفهوم البحث عن البيانات و استرجاعها، فمن وضعية تسعى للعثور على المعلومة صار الهدف من هذه العمليات الفرز و الانتقاء. أمام هذه التغيرات، تطورت وسائل البحث والاسترجاع لتكتسي أبعاداً أكثر تخصصاً و ذكاءً بوساطة ما أفرزه الإطار الجديد من آليات و برمجيات متعددة تسعى لمزيد من التمكن من المعلومات و طرق معالجتها. لكن أمام تفاقم حجم البيانات و تشعب متطلبات المستخدمين ظهرت على شبكة الإنترنت خدمات أخرى توفر قيمة إضافية للمعلومة المسترجعة، هدفها إعانة المستخدم على توفير أقصى قدر ممكن من المصداقية و الجدوى العلمية للتصرف و أخذ القرار .

المقدمة

لقد صار من البديهي اليوم أن نشاهد معظم المستخدمين على شبكة الإنترنت بمختلف أصنافهم يعتمدون أدوات البحث و الاسترجاع "التقليدية" للمعلومات، تلك الطرق و الأساليب التي كانت بالأمس القريب نموذجاً أحدث ثورة في عالم نظم المعلومات و إدارتها. فقد كانت غالبية النظم تقتصر على المعلومات المرجعية (Référentielles) التي تشير إلى وجود الكتب و المقالات دون إتاحة الإمكانية للوصول إلى محتوياتها المعرفية. وقد شكلت في هذا المجال قواعد البيانات للمكتبات الجامعية (OPAC) النموذج الأكثر شيوعاً في مجال البحث و الاسترجاع حتى إن بعض المؤسسات التجارية توختها كطريقة

للتجارة و الربح (Dialog, Questel). وتأتي اليوم البدائل التكنولوجية الحديثة لتغيير كل الأسس العملية لهذه الأنظمة ليحل النص الكامل (Full Text) محل الوصفة الببليوغرافية، وتحل المعلومات متعددة الوسائط محل المعلومة النصية و تستبدل الأجهزة المكتبية المنعزلة و الشبكات المحلية بأنظمة أكثر شيوعاً وانفتاحاً كالشبكات القطرية و الإقليمية و الدولية المبنية على برمجيات متطورة و تصميمات مفتوحة و موزعة. وتمثل شبكة الإنترنت اليوم الإطار العام الذي نبني حوله كل نظم المعلومات بجميع مكوناتها من بنية للشبكية و هيكله للمحتويات المعرفية و نوعية للخدمات الاتصالية داخلها.

فدون الرجوع إلى الإحصائيات التي يزخر بها عدد كبير من الكتب و المراجع حول تطور شبكة الإنترنت من حيث عدد الأجهزة المرتبطة و المستفيدين من كل خدماتها، و كذلك العدد الهائل لمحتوياتها، فإن عصر المعلومات اليوم يتميز بنقله نوعية و قفزة كمية هائلة من حيث حجم الوثائق المتوفرة على الشبكة و تنوع محتوياتها. هذا الحجم الهائل و المتطور يومياً إن لم نقل باستمرارية مذهلة، غير مفهوم البحث و الاسترجاع من ظاهرة البحث و الوصول إلى المعلومة ليضع المستخدم في حالة انتقاء و اختيار للمعلومة الأكثر جدوى و نفعاً لأخذ القرار أو البدء في إنجاز العمل. فقد صار اليوم ممكناً لأي كان، مؤسسة أم فرداً، تحميل بياناته و جعلها على ذمة كل قارئ حقيقي أو افتراضي على الشبكة، الشيء الذي أعطى المعلومات على الشبكة شمولية أكثر للمعرفة الإنسانية من ناحية المحتوى لكن قلص من جدوى فعاليتها نظراً لوجود الغث و السمين منها جنباً إلى جنب و تزامن المحدث منها و القديم دون تنبيه و إتلاف بعضها أو تحديثه دون سابق إعلام.

كل هذه المعطيات و غيرها من الخصائص الإضافية التي أفرزتها التقنيات الحديثة للمعلومات و طرق استخدامها المتطورة جعلت من الإطار العام للبحث و الاسترجاع للمعلومات عنصراً دائماً الحركية و التفاعل مع متغيرات العناصر الأخرى المرتبطة به كشبكات الاتصال عن بعد و البحوث في مجالات علمية كاللسانيات و علم المكتبات و الذكاء الاصطناعي و الترجمة. وقد شكلت هذه التغيرات عنصراً هاماً في إنجاز و تطوير الآليات و الطرق الحديثة لمعالجة الوثائق الإلكترونية بقصد استرجاعها بأكثر نجاعة و تفادياً لما عرف تقليداً بالصمت و الضجيج البحثيين اللذين أضيفت لهما اليوم عناصر تقييمية أخرى كالشمولية و الجدوى العلمية. و بمرور الزمن، صارت محركات البحث بجميع أصنافها آليات عادية بالنسبة للمستفيدين يتم استخدامها دون تكوين مسبق أو

الإنترنت والمكتبات.. مقاربات أولية

دراسة معمقة ضرورية، كما أن لمفعول المنافسة بين أصحاب المحركات انعكاسات جعلت واجهات البحث اليوم تتسم بتفاعلية قصوى مع المستفيد و بخدمات إضافية تسهل عمليات التوجيه لمصادر المعلومات و تساعد على انتقاء الأجدى منها .

إلا أن مجال البحث عن البيانات بقي كسائر مكونات عالم المعلومات و تقنياتها دائم النمو والتطور و الحركية . و قد ظلت أدواته حية تتطور معتمدة في ذلك إنجازات هندسية الذكاء الاصطناعي و متأثرة بمنهجيته في معالجة المسائل الصعبة ذات الحلول المتشعبة . لذا فإن وصفنا في بداية هذا البحث لمحركات البحث بأنها تقليدية نابع من مقارنتها بنظيراتها الحديثة الأكثر ذكاء و مردودية بالنسبة للمستفيد . فما يعرف اليوم بالأعوان الذكية تشكل نموذجاً جديداً من أدوات البحث يجدر التوقف عندها لتحليل طرق عملها و كيفية التعامل معها، تأكيداً لما توفره من ربح في الجودة و ضماناً للقيمة المضافة التي يسعى إليها كل باحث للمعلومات على الشبكة.

انطلاقاً من هذه المعطيات الأولية سنتطرق إلى العناصر الثلاثة التالية :

- طرق عمل أدوات البحث الكلاسيكية

- طريقة عمل الأعوان الذكية

- أشكال خدمات القيمة المضافة للمعلومات

1 . أدوات البحث وطرق عملها في الكشف و البحث و الاسترجاع للمعلومات:

قبل الوصول إلى الحديث عن جانب الذكاء لدى آليات البحث الحديثة، يجدر بنا التعرض إلى طرق عمل أدوات البحث التقليدية على الشبكة و كيفية معالجتها للبيانات من الناحية التطبيقية، ذلك الجانب الذي يخفى على أغلبية المستعملين نظراً لحدوثه في عمق النظام (Back Office) بعيداً عن الواجهات التفاعلية مع المستفيدين (Front Office).

1 . 1 مكونات أدوات البحث على شبكة الإنترنت:

تتكون أدوات البحث أساساً من الأجزاء الثلاثة التالية التي تكمل بعضها:

1 . 1 . 1 . البرنامج الآلي (Robot): هو عبارة عن برنامج يقوم " باصطياد " كل جديد أو محدث من صفحات الويب وفق رقعة اتساع جغرافية، لغوية أو موضوعية يتم ضبطها مسبقاً من قبل مصمم المحرك . كما يمكن لمصممي مواقع الويب استدعاء البرنامج الآلي لأي محرك للقيام بعمل الفهرسة للصفحات المكونة لموقعه، كما يمكنه أيضاً منع ذلك

حرصاً على سرية الموقع أو خصوصيته. وتتم هذه العملية في ترددية يضبطها صاحب المحرك تتفاوت من الدقيقة إلى اليوم والأسبوع والشهر.

1. 1. 2. المفهرس (Indexer): وهو أيضاً برنامج يقوم بفهرسة كل الوثائق التي قام البرنامج الآلي باستيراد نسخ منها. تتم هذه العملية بمجرد كل محتوى الوثيقة لاستخراج محتواها من مصطلحات و تعابير يتم اختيارها وفق ترتيب معين للأهمية التي تكتسبها داخل الوثيقة. ويعبر فعلاً هذا البرنامج المفهرس أولوية تنازلية لأماكن وجود المصطلحات يكون العنوان الإلكتروني للوثيقة أهمها (URL) تليه في ذلك بيانات التعريف (Metadata) على مستوى رأس الوثيقة ثم العناوين و العناوين الفرعية فبقية النص من بدايته إلى نهايته. وتضاف كل هذه المصطلحات داخل كشاف موحد يكون منطلقاً لعمليات البحث.

1. 1. 3. واجهة البحث (Interface): هي الجانب المرئي من المحرك وهي الواجهة التي تخول المستخدمين تحديد موضوعات البحث لديهم. ويتم ذلك إما بإدخال تركيبات البحث مباشرة إن كان الأمر يتعلق بمحركات البحث أو باتباع هرمية موضوعية إن تم اللجوء إلى أدلة البحث الموضوعية. في الحالة الأولى يقوم النظام بالتثبت من وجود المصطلحات المستعملة داخل المفهرس ثم يتم عرض النتيجة في شكل قائمة لعناوين الوثائق (URL) التي وقع ضبطها كمصادر تحتوي على مفردات البحث المستعملة. أما بالنسبة للبحث عبر الأدلة فيتم اتباع تدرج المستخدم مع إعطائه كل مرة جملة من عناوين الوثائق التي تماشي مستوى الغمق الذي بلغه في الهرمية.

كما تجدر الإشارة أيضاً إلى أن هذه الطريقة في التكشيف والبحث مأخوذة من نمطية نظم معالجة قواعد البيانات و إدارتها التي تبني على مبدأ التكامل المذكور بين رصيد الوثائق المخزونة تسلسلياً و المفهرس الكشاف الذي يتم من خلاله إعادة تهيئة محتوى الوثائق المجمعّة باستخراج كل المصطلحات و سردها هجائياً في ملف مقلوب يعتمد أولوية الرموز فالأرقام ثم الحروف. و بالتالي يكون السرد الهجائي للمفهرس (اعتماد الموضوعات كمدخل رئيسية) مكملاً للترتيب التسلسلي للملف الأستاذ أو قاعدة الوثائق (اعتماد الوثيقة كوحدة تسلسلية) في عملية البحث (بالموضوع) و الاسترجاع (بالجذاعة أو الوثيقة الكاملة). لهذا السبب توصف شبكة الإنترنت عادة بأكثر قاعدة بيانات كونية .

1. 2. أصناف أدوات البحث و طرق عملها:

للمزيد من التعرف على أصناف آليات البحث و طرق عملها نورد فيما يلي وصفاً إضافياً

الإنترنت والمكتبات. مقاربات أولية

للعمليات الرئيسية لبناء نظم المعلومات الآلية على الشبكة و كيفية استخدامها. وسنتولى خلافا لما هو متوقع من عرض تعريفي لأدوات البحث، التعرض لجانبها العملي من حيث تطابقها مع العمل التوثيقي المعهود لدى الاختصاصيين من فهرسة و تكشيف ثم بحث واسترجاع للمعلومات.

1. 2. 1. الفهرسة و التكشيف:

كما أوردناه سابقا، تمثل عملية التكشيف المنطلق الرئيسي لتكوين واحدة من أهم عناصر أنظمة البحث الآلية وهي الفهارس الموضوعية. وتتم هذه العملية بطريقة مطابقة لنظم إدارة قواعد البيانات إذ يتم استخراج كل المصطلحات من الوثائق المجمعّة باستثناء قائمة المصطلحات الجوفاء التي يحددها مصمم المحرك لتفادي شحن الفهارس و إثقالها بكلمات لا تُستعمل بتاتا في عمليات البحث. ثم يتم إدراج نتيجة الفهرسة بالكشاف مصحوبة بعناوين الوثائق التي تم استخراج المصطلحات منها وذلك في ترتيب هجائي يعتمد أولوية الرموز ثم الأرقام فالحروف.

انطلاقا من هذه الخصوصيات للعمل الداخلي والخارجي لأدوات البحث بالشبكة، عمدت بعض المؤسسات المختصة إلى إضافة بعض المميزات الخاصة بها فسارعت بدعم هذا العمل التكشيفي الآلي بتعيين قدرات بشرية مختصة تعمل على تصنيف الوثائق المجمعّة وعرضها في قوائم موضوعية هرمية. وقد عرفت هذه الأدوات بالأدلة الموضوعية مقارنة بالمحركات رغم أن كلا الصنفين يتشابهان إلى درجة أنه يصعب التمييز بينهما نظرا لتوخيها كلتا الطريقتين (فهرسة آلية و توضيب بشري) في الوقت نفسه. أما الفوارق بين الوسيطتين فتتلخص في مدى عمق كل جانب من جانبي التكشيف و الفهرسة آلية كانت بالنسبة للمحركات أو بشرية بالنسبة للأدلة الموضوعية. فمن الخصائص المميزة للأدلة مثلا أنها تبني فهارسها الآلية انطلاقا من الرصيد الوثائقي المعتمد بالدليل فقط بينما تتعدى المحركات الجزء البسيط من الدليل الذي توفره لتقوم بتكشيف و فهرسة أكبر عدد ممكن من الوثائق التي يقوم البرنامج الآلي بحصرها واستيرادها أثناء جولاته عبر قنوات الإنترنت، والتي لا يقع بالضرورة اعتمادها كليا في تهيئة الدليل الموضوعي الذي تقترحه. ويعد دليلياهو (Yahoo) من أشهر الأدلة الموضوعية على الإنترنت نظرا لعمق هرمية الترتيب الموضوعي الذي توفره واجهة البحث. كما يعد محرك ألتافستا (Altavista) من أهم المحركات عمقا وشمولية.

1. 2. 2. خدمات البحث:

أما فيما يخص عمليات البحث، فلكل أداة واجهتها المميزة وخدماتها المتطورة، رغم أن معظمها يتشابه من حيث التقنيات المعتمدة، والتي يمكن تلخيصها فيما يلي :

1 . استخدام المنطق البوليني بأدواته الثلاثة (و، أو، ما عدى) و لو اختلفت أشكال كتابتها. فبعضها يعتمد الرموز الرياضية (+ - *)، أو المحارف (AND ,OR, NOT)، أو اختيار ذلك في قائمة اقتراحات (معروضة أو منسدلة).

2 . استخدام اللغة الطبيعية. تعتمد بعض المحركات هذه الطريقة لتسهيل عملية البحث للمستخدم العام الذي لا يُطلب منه التمكن من أدوات الربط البولينية علماً وأنه يقع إعادة تفكيك السؤال من قبل النظام في مستوى ثان من الواجهة لمعالجته وفق المعايير البولينية وبحضور ملفات الكلمات الجوفاء التي تُستخدم لإزالة ما ورد منها بتركيبة البحث الطبيعية.

3 . استخدام أدوات الجوار والتقارب بين المصطلحات لتحديد أدقّ لنوعية الوثائق المطلوبة. تكون هذه الأدوات تقنية هامة جداً بالنسبة لعمليات البحث للوثائق النصية إذ هي تمكن من حصر نتيجة البحث على الوثائق التي تكون فيها المصطلحات المطلوبة متقاربة وفق المسافات أو عدد الكلمات التي تفصل بينها.

4 . توجيه المستفيد إلى مصادر خاصة للمعلومات كمواقع الويب، أرشيفات الأخبار (News)، الصفحات الصفراء. توفر المحركات للمستفيد إمكانية الاختيار لمجالات البحث بتصويب عملياته البحثية تجاه نوعية معينة من الوثائق تستهدف، علاوة على ما وقع ذكره، لغة معينة أو جهة محددة.

5 . التركيز على نمط معين من البيانات متعددة الوسائط. توفر بعض أدوات البحث إمكانية التخصيص في استهداف نمطية معينة من الوثائق المتعددة الوسائط كالصور أو الفيديو أو الملفات الغنائية.

ولإن مثلت الأدلة الموضوعية والمحركات أهم الأدوات البحثية على الشبكة، فإن تفاقم حجم المعلومات من جهة ومحدودية مساحات التكشيف لهذه الأدوات من جهة أخرى جعلت عمليات البحث منقوصة دائماً نظراً لعدم شمولية فهارس المحركات و تباين عمق الأدلة الموضوعية. لذا، كان من الضروري عند كل عملية بحث التوجه إلى أكثر عدد ممكن من المحركات و الأدلة لاسترجاع أكثر عدد ممكن من الوثائق. إلا أنه صار من غير الممكن

الإنترنت والمكتبات.. مقاربات أولية

القيام بهذه العمليات لأسباب عدة منها مدة البحث المطولة التي تنتج عن هذه العمليات المتكررة، ازدواجية الوثائق المسترجعة من قبل عدد كبير من المحركات، اختلاف لغات التساؤل المستعملة، إلخ.

لهذه الأسباب وغيرها ظهرت شاكلة أخرى من الأدوات عرفت بالمحركات المتعددة تختلف عن سابقتها من حيث الوظيفة و التركيبية الداخلية. فمن خصائص هذه المحركات المتعددة أنها في حقيقة الأمر لا تطور وحدها عمليات التوريد والتكشيف كبقية المحركات، بل هي في نهاية الأمر واجهات تطبيقية تقوم بدور الوسيط بين المستفيد و جملة من المحركات والأدلة. يأخذ المحرك المتعدد تركيبة البحث ليتولى إرسالها إلى مجموعة من أدوات البحث التي صُمم للتعامل معها. تعالج هذه الأدوات التركيبية انطلاقاً من فهارسها ثم تتولى إرسال النتائج للمحرك المتعدد الذي يقوم عادةً بإعادة هيكلتها لإزالة ازدواجية الوثائق و إضفاء لونه الخاص عليها مع ذكر مصدر كل وثيقة تم العثور عليها. من أشهر هذه المحركات المتعددة نذكر على سبيل المثال Alltheweb، Use-It، Mamma، Metacrawler.

1. 2. 3. عملية الاسترجاع:

أما استرجاع نتائج البحث فيكون عادة في شكل قائمة لمجموع الوثائق التي تستجيب لعناصر البحث و التي وقع حصرها خلال عملية التكشيف. وتكون هذه القائمة في شكل ملخصات تذكر عنوان الوثيقة على الشبكة (URL)، ملخص لمحتواها، حجمها... مع إضافة الربط التشعبي الذي يوصل إلى موقعها على الشبكة. كما تمكن بعض المحركات المستفيد من حصر مدة البحث، تحديد عدد الوثائق المسترجعة في الشاشة الواحدة، و كذلك تركيبة التسجيلية المعرفة للوثيقة.

أما أولوية سرد النتائج فهي تخضع أيضاً إلى معايير تقييمية يتم بوساطتها إسناد قيمة ترتيبية للوثيقة وفق تطابق محتواها مع تركيبة البحث. تسند هذه القيم عادة وفق مقاييس معينة كنسبة تكرار المصطلح في الوثيقة، مكان وجوده فيها (الرأس، بداية الوثيقة، نهايتها)، قيمة مكان وجودها في الوثيقة (عنوان رئيسي، عنوان فرعي، تذييل...)، تاريخ إنشاء الوثيقة أو تاريخ تحديثها. وتوجد اليوم منافسات شديدة بين مصممي مواقع الويب لجعل أعمالهم تتصدر نتائج البحث ضمن الوثائق العشر الأولى ضماناً لقراءتها أو تثبيتها لقيمة وظيفتها الإشهارية. فقد يعتمد بعضهم إلى شحن رؤوس و ثنائقهم بأكثر عدد ممكن من المصطلحات مع الإطناب في تكرارها (Spamming) التي تقوم المحركات بإعطائها القيمة

القصوى .غير أن العديد من المحركات تفتنت إلى هذه الطرق الملتوية وواجهتها بعدم فهرسة الصفحات التي يلاحظ بها اعتماد هذه الطريقة (Anti Spamming).

2 . الأعوان الذكية في البحث عن المعلومات واسترجاعها:

رغم تنوع وسائل البحث والاسترجاع للمعلومات وتنمية قدراتها على الاستجابة لطلبات المستفيدين البحثية، يبقى تدفق البيانات أقوى وأسرع من قدرات أي وسيلة تمكن من حصره والتحكم فيه .لذا تشهد اليوم أدوات البحث والاسترجاع للمعلومات اعتمادا متزايدا على إنجازات هندسية الذكاء الاصطناعي والنظم الخبيرة لتقدم أفضل الخدمات للمستفيدين . فما هي إذا خصوصيات هذا الذكاء وهذه الخبرة؟ وما هي طرق إدماجها في المنظومات البحثية على الشبكة بصفة عامة و الأعوان الذكية بصفة خاصة ؟

2 . 1 . الذكاء الصناعي :

رغم أن عبارة "الذكاء الصناعي" قد ظهرت بطريقة رسمية لأول مرة سنة 1969 خلال المؤتمر الأول حول الذكاء الصناعي في واشنطن، يبقى جون ماك كارثي (John McCarthy) أول من استخدم هذه العبارة سنة 1956 في معهد التكنولوجيا في ولاية ماستشوسستس الأمريكية (MIT) أما عن المفاهيم التي يعرف بها الذكاء الصناعي فهي تنقسم عادة إلى شقين أولاهما يعنى بإعادة تمثيل الذكاء البشري عن طريق الحاسوب الآلي، و ثانيهما يهتم بتوسيع رقعة استعمالات هذا الأخير بإعطائه قدرة معينة من التخمين و التصرف المستقل في مجالات معرفية متعددة كالفيزياء و الرياضيات و الهندسة و الفلسفة و علم النفس .

أما أشكال استخدام الذكاء الصناعي فهي متعددة نذكر من أهمها :

2 . 1 . 1 . الأنظمة الخبيرة:

تعددت مفاهيم الأنظمة الخبيرة عبر مختلف مراحل تطورها لكن يمكن تلخيص ذلك بالقول: إن هذه الأنظمة تكون جملة من البرمجيات تتصرف آليا في مجموعة من المعطيات المختصة لتولد تصرفا خبيرا يقترب كثيرا من تصرف الإنسان في معالجة المسائل المعقدة بإضافة إلى واجهاتها الذكية و محركاتها "الاستنتاجية" (Moteur d'inférence) "التي تمكنها من استخلاص النتائج بمقارنة البيانات المخزنة والحقائق المعرفية الثابتة، تبني هذه الأنظمة كامل تصرفاتها انطلاقا من قاعدة معرفية رئيسية تنظم كل الفرضيات الممكنة وفق سياق عملية البحث الجارية.

الإنترنت والمكتبات.. مقاربات أولية

2. 1. 2. الشبكات العصبية:

هي جملة من الإجراءات الأساسية التي تمكن عند إدماجها فيما بينها من القيام بعمليات معينة يطلبها المستفيد عند الحاجة. لذلك من الضروري للمستخدمين من خدمات الشبكات العصبية القيام بعمليات أولية لتدريب الشبكة تدريجيا حتى تتمكن في النهاية من الاستجابة لكل مطالبهم. وتتم هذه التدريبات في شكل تلقينات متكررة للتصرفات المطلوبة من قبل الشبكة من خلال برامج مختصة تتولى السهر على ضمان تطابق تصرفات الشبكة مع متطلبات صاحبها.

2. 1. 3. المنطق المعتم:

تتقاطع إضافات الشبكة العصبية مع تقنيات المنطق المعتم (Logique flou) الذي يهدف إلى بناء الاستنتاجات والقرارات النهائية انطلاقا من معطيات عامة و غير ثابتة بوساطة عدد كبير من البرمجيات المندمجة والمركبة. وضع نظرية المنطق المعتم " لفي زادة " في جامعة واشنطن خلال الستينيات انطلاقا من مقارنته باشتغال الذكاء البشري في استعمال البيانات التقريبية و غير الثابتة للوصول إلى الحلول و أخذ القرار. فبالمقارنة مع المتطلبات الدقيقة للبرمجيات الإعلامية التقليدية، صار من الممكن اتخاذ منهج الفرضيات لمعالجة المسائل الصعبة و اتخاذ القرارات المعقدة.

كما تتبين علاوة على ذلك أشكال استخدام الذكاء الصناعي من خلال الأعران الذكية والروبوتية و الألعاب الإلكترونية و برامج المعالجة الآلية للغات الطبيعية. ولعله في هذا المجال يجدر بنا التوقف قليلا عند ظاهرة الأعران الذكية لاسترجاع المعلومات و معالجتها قبل التعرض إلى خدمات القيمة المضافة لديها.

2. 2. خصائص الأعران الذكية :

لعل أبرز خدمة تقوم بها الأعران الذكية تتمثل في حفظ هوية و خصائص مستعملها للاستجابة لخدماتهم آليا دون تدخل مسبق منهم. فبعد تسجيل كل المعطيات حول المستفيد تتولى الأعران الذكية القيام بعمليات البحث و الانتقاء ثم إيداع النتائج الأخيرة في علبة رسائله الإلكترونية أو توجيهه آليا إلى المواقع التي تحتوي على البيانات المطلوبة. ولئن كانت هذه طريقة عمل الأعران الذكية في البحث عن المعلومات المتبعة خصوصا من قبل نظم الدفع الآلي للمعلومات (Push Technologie)، فلأعران الذكية عدة وظائف أخرى تمتاز بها كالقيام بالأعمال الروتينية التي يسعى المستفيد التخلص منها كالإجابة عن رسائله الإلكترونية،

القيام بعمليات الحجز الآلي، التصرف الموقوت في التنبيه و التذكير و أخذ المواعيد. تتكون الأعوان الذكية من عدة مكونات تساهم في إضفاء صفة الذكاء عليها. فهي في حقيقة الأمر جملة من البرمجيات والخوارزميات المندمجة تتداول في إنجاز مختلف الوظائف الخاصة به و تتداخل على مستويات عدة لإتمام عملية معينة.

يمتلك العون الذكي عادة قاعدة معرفية مسبقة يتصرف على ضوءها للاستجابة لطلبات المستفيد. على أنه يمكن للعون الذكي أن يطور هذه القاعدة المعرفية بمعارف أخرى يمكنه اكتسابها خلال عمليات التعلم التي يخضع إليها طوال مدة عمله .

كما للعون الذكي أيضا روتينيات آلية لتصميم نوعية و طريقة الخدمة التي يستوجب عليه إدلائها للمستفيد. وفي عندما يتعذر عليه القيام بذلك يمكنه إما التفاعل مع المستفيد أو اللجوء إلى أعوان آخرين للترؤد بمزيد من الإرشادات بقصد الإجابة عن تساؤل أو القيام بعملية محددة.

للقيام بكل هذه الخدمات، تمتاز الأعوان الذكية بالخصائص التالية :

- الاستقلالية الذاتية في التصرف و أخذ القرار و المتابعة المحكمة لكل العمليات.
 - التعاون و العمل المشترك مع أعوان ذكية أخرى أو التفاعل مع المستفيد لاستجابة أمثل لطلباته.
 - القدرة على التدريب و التكيف مع حاجيات المستفيد بالملاحظة المستمرة للخدمات السابقة.
 - المعالجة الموجهة للأشياء من خلال تعبيرات البحث للمستفيد.
 - التدريب الذاتي و ذلك بالتكيف مع تغيرات خصائص المستفيد التي هي بصدد خدمته.
- 2 . 3 . درجات الذكاء لدى الأعوان الذكية في البحث عن المعلومات :

كامتداد لتطورات أدوات البحث التي ذكرناها سابقا، وقع إدراج الأعوان الذكية كجزء أساسي لدى عدد كبير من محركات البحث على شبكة الإنترنت. لذا نجد اليوم عددا هائلا من أدوات البحث الذكي التي تقدم أنواعا مختلفة من الخدمات المتفاوتة الذكاء من الإعانة عن البحث إلى التكفل التام بعمليات البحث والاسترجاع ثم البث و التوزيع. و سنتعرض هنا إلى بعض النماذج من هذه الأدوات والوسائل وفق درجة ذكائها و قدرتها على التعامل المتطور مع المستفيد. فعادة ما يتم تصنيف الأدوات الذكية وفق الأجيال التي تنتمي إليها. إلا أننا سنتوخى في هذا البحث تقسيماً يماشى محدودية أو شمولية العمل الذي تقوم به.

2 . 3 . 1 . أدوات البحث ذات الذكاء النسبي:

في حقيقة الأمر وقعت إضافة جانب كبير من الذكاء على كل محركات البحث التقليدية وذلك تحت وطأة المنافسة التي تشهدها كل الخدمات على شبكة الإنترنت. فالخدمات الإضافية من نوع تحديد مجالات الإشهار التي توفرها المحركات للمستفيدين وفق طلباتهم أو خدمات تشخيص البرقيات الإخبارية التي توفرها محركات و أدلة كـ"ياهو" و "إكساييت" من خلال صفحاتهم المشخصة (MyYahoo - MyExcite)، كلها خدمات يمكن وصفها بالذكاء المحدود إذا تمت مقارنتها بالذكاء الكامل الذي سنتعرض إليه لاحقاً.

على أن هذا الجانب من الذكاء تم تطويره و إعطاؤه البعد اللازم حتى يشكل بديلاً كاملاً للمحركات والأدلة الحالية، و ذلك في شكل أدوات إضافية بحثية تلقى اليوم إقبالاً كبيراً في مجالات البحوث المختصة و الأكاديمية.

هذه الأدوات البحثية ذات الذكاء النسبي أو المحدود تأخذ من هندسة الذكاء الصناعي قدرتها على التعامل مع المستفيد بطريقة آلية و متكررة دون ضرورة تحميلها أو الالتجاء إليها كلما وقعت الحاجة لها.

فمحرك البحث "إنفورمانت Informant2" على سبيل المثال وشريكه الجديد "تريسرلوك" Tracerlock)3) يمكنان المستخدم من التسجيل لديهم ثم تحديد مجالات بحثه. و يتكفل المحرك بعد ذلك بالقيام بالبحث المتواصل على المواقع و الوثائق التي تتعلق بموضوعات البحث المذكورة ليرسل بقائمة فيها على العنوان البريدي للمستخدم أو ليعرضها عليه وفق حدائتها عند قدوم هذا الأخير على موقع المحرك .

تقوم محركات أخرى من هذا النوع بعمليات المراقبة المستمرة على المواقع لترصد التغيرات أو التحديثات التي تتم عليها لتبلغ بذلك كل مستفيد مسجل لديها عبر عن اهتمامه بها (مثال محرك ناتمايند4) .

تبرز ظاهرة الذكاء لدى محرك "آسك دجيفز5 Askjeeves) "باعتماده مبدأ النظم الخبيرة في معالجة اللغات الطبيعية للمستخدم و إعادة صياغتها في تركيبات بحثية تتم معالجتها بواسطة خوارزميات أسنوية لاستخراج الكلمات المفاتيح منها. يقع اعتماد هذه الكلمات إثر ذلك في محاورة النظام الخبير المعتمد عليه إما لاسترجاع عناوين الوثائق التي تماشي سؤال المستخدم أو لإعطاء جملة من المقترحات البديلة لكيفية طرح السؤال المتعلق بموضوع البحث الجاري. فعادة ما يخزن هذا المحرك الأسئلة المطروحة عليه من قبل كل

المستفيدين و يقوم باقتراحها كبداية للسؤال كلما دعت الحاجة إلى ذلك .
تطورت هذه النظم و أفرزت فصيلة أخرى من النظم المتطورة ذات قدرات أكبر و أشمل.

2 . 3 . 2 . أدوات البحث ذات الذكاء المتطور:

يعد هذا النوع من النظم الذكية آخر ما وصلت إليه تقنيات البحث و الاسترجاع للمعلومات على شبكة الإنترنت .وهي تندرج في مجال عرف كثيرا بمجال "تكنولوجيا الدفع" (PushTechnologie) نظرا لاعتماده طريقة معاكسة للنظم التقليدية التي تستوجب من المستفيد السعي الدائم و المتكرر إليها للحصول على الخدمة أو المعلومة.

فالنظم الذكية المتطورة اليوم هي التي تسعى من ناحيتها للوصول إلى المستفيد لتقديم الخدمات إليه بطريقة آلية و غالبا ما تكون مجانية دون طلب منه إذ يكفي هذا الأخير التعريف بنفسه لدى المحرك الذكي أو الموزع الذي يأويه ليتصل باستمرار بآخر ما صدر في مجال اختصاصه من معلومات و أخبار جديدة.

أما الجديد في هذا المجال فهو الطريقة التي تتم بوساطتها عملية الاتصال بين المحرك و المستفيد .هذه العملية مأخوذة عن تقنية قديمة عرفت بها المكتبات باسم النشر الانتقائي للمعلومات إذ يتم إيصال المعلومات الواردة حديثا على المكتبة أو مركز التوثيق عبر البريد أو مباشرة إلى أصحابها وفق قائمة يقع ضبطها مسبقا وفق مراكز اهتماماتهم. مطابقة لذلك يتم من خلال شبكة الإنترنت إرسال المعلومات في كل أشكالها و بطريقة مسترسلة و دائمة إلى المستفيدين المسجلين الذين يتلقون تلك المعلومات في شكل قوائم منسدة أو أشرطة انسيابية تظهر على شاشاتهم أثناء عملهم على الإنترنت .

من أهم الموزعات التي عرفت بهذه الخدمات الذكية، نذكر على سبيل الذكر و لا الحصر موزع "بوينتكاست (PoinsCast) " و "ماريمبا (Marimba) " و "إنفوقايت (Infogate) " وغيرها من الموزعات في العالم التي تستعمل تكنولوجيا الدفع الانتقائي للمعلومات.

يبقى في هذا المجال أن نذكر أن كل هذه الأشكال من الخدمات بكل درجات ذكائها تبقى جامدة و غير متحركة على مستوى محتوى المعلومة الموصولة إلى المستفيد .فهذه الأنظمة تعنى أساسا بعملية الاتصال و لا تتصرف على مستوى المحتوى المعلوماتي للوثيقة أو الخدمة المنجزة .لهذا يبقى مجال الخدمات المعلوماتية في البحث و الاسترجاع منقوصاً وفق نظرنا من شيء نعدده أساسيا و جوهريا ألا وهو مسألة الخدمات ذات القيمة المضافة للمعلومات .لهذا الغرض نخصص الجزء الموالي من هذا البحث إلى هذا الموضوع الهام.

3. خدمات القيمة المضافة للمعلومات :

نهدف من خلال مفهوم الإضافية في هذا المجال الدلالة على كل ما يحدث للمعلومة المسترجعة من وظائف تكميلية تسهل درجة و كيفية استخدامها . و هذه الخدمات في حقيقة الأمر تهم المستفيد القارئ أو المستفيد المصمم للمواقع والوثائق الإلكترونية على حد سواء . ونعني بذلك مثلا خدمات الترجمة متعددة اللغات و التقييم الموضوعي وفق تركيبة البحث و كذلك عنصر الإيقاع للوثيقة في مجالها المعرفي في الشبكة .

3 . 1 . الترجمة الآلية:

تمثل خدمة الترجمة الآلية للوثائق الإلكترونية على الشبكة من الخدمات المضافة الأكثر رواجاً اليوم . فأغلبية المحركات و المواقع المختصة توفر إمكانية الترجمة الآنية للوثائق المسترجعة في عدد كبير من اللغات . نذكر من جملة هذه المحركات محرك Voilà الفرنسي (http://www.voilà.fr) ، Altavista ، (http://www.altavista.com)...

لكن تجدر الإشارة إلى أن هذه الترجمة تبقى إلى حد اليوم ترجمة أولية و غير معمقة نظراً لأسباب عدة من أهمها :

* عدم اعتماد المحركات في هذه العملية الخوارزميات اللسانية المختصة للمعالجة الموضوعية و النحوية للنصوص بل تقتصر على تقنيات قوائم المصطلحات الموازية مع خوارزميات نحوية مبسطة للقيام بالتراجم الموازية دون تعمق في التحليل الموضوعي والمعرفي للوثائق .

* سطحية المجال المعرفي للمحركات ، وبالتالي عدم قدرتها على الترجمة المختصة في مجالات معرفية دقيقة تستوجب مجموعة من المصطلحات الخاصة .

3 . 2 . التحليل الموضوعي:

ظهرت خلال السنوات الأخيرة نوعية جديدة من التطبيقات المختصة هدفها مواجهة التدفق المتصاعد للمعلومات على شبكة الإنترنت ، هذا التدفق الذي كما ذكرناه سابقاً غير مفهوم البحث عن المعلومات من وضعية بحثية إلى عملية انتقائية نظراً للقدر الهائل من المعلومات التي تسترجعها المحركات مع كل عملية بحث . وظيفة هذه البرامج ذات القيمة المضافة إعادة النظر في محتوى نتائج البحث لصياغتها من جديد وفق أهميتها مع تركيبة البحث الأولية و تقديمها للمستفيد في شكل أكثر وضوحاً و شفافية .

من هذه البرمجيات المختصة ذات القيمة المضافة نذكر مثلاً برنامج UmapWeb

(http://www.umap.com) الشهير الذي يعالج نتائج البحث التي يقوم بها بنفسه على الشبكة أو التي يتم تزويده بها ثم يقوم بعرضها بطريقة خرائطية و في أشكال مختلفة لإبراز مدى علاقة الوثائق المسترجعة مع نواة السؤال المطروح و مدى ترابطها فيما بينها.

3.3. عنصر الإيقاع على الشبكة (Web Impact Factor):

من خدمات القيمة المضافة الأخرى التي بدأت تأخذ مكانها خصوصا بالنسبة للباحثين والمؤسسات العلمية المختصة، نذكر عنصر الإيقاع المعرفي لأي وثيقة على الشبكة. و نعني بذلك مدى إشعاع الوثيقة على محيطها العلمي والمعلوماتي و ذلك بالنظر إلى درجة اعتمادها كمرجع يقع الاستناد إليه و بالتالي ذكره أو إنشاء روابط تشعبية نحوه. فمن شأن هذه الخدمة المضافة أن تضفي طابعا علميا على محتوى الوثيقة و أن تعطيه أكثر مصداقية لدى القارئ.

اعتنت بهذه الظاهرة عدة محركات على الشبكة و أدرجتها كخدمة من جملة الخدمات الإضافية التي تقترحها على مستوى واجهات بحثها الرئيسية. مثلنا في ذلك محركات Altavista و Google كما يبينه الرسم التالي.

و يشكل هذا العنصر اليوم أحد الركائز الأساسية لبرامج مختصة تعنى بإشكاليات القيمة العلمية للمعلومات على شبكة الإنترنت و طرق الإقرار بمصداقيتها. فعلاوة عن المعايير التي يذكرها Wilfred Lancaster في كتابه "نظم استرجاع المعلومات 6" لتقييم خدمات المعلومات (التكلفة و الوقت و النوعية)، يشكل تقييم المحتوى عنصرا لا يمكن تجاهله. و لعل أبرز النماذج التي تعطي هذا العنصر أهميته القصوى اليوم هو نموذج Netscoring7 المتخصص في مجال المعلومات الطبية لفرط ما يكتسبه هذا المجال من حساسية المحتوى و ضرورة التثبت من مصداقيتها.

الخاتمة :

لئن كان الحافز الرئيسي للمتطرق لهذا الموضوع نابعا من تعامل يومي مع أدوات البحث عن المعلومات في شبكة الإنترنت، فإن ما جعل المسألة تأخذ شيئا من الخطورة هو ما وقع الإقرار به من أن أغلبية المستفيدين سواء كانوا من المستعملين العاديين - و ليس في ذلك غرابة - أو من المهنيين، وهنا تكمن المشكلة، غير مدركين لأبعاد هذه الخدمات حتى إنهم غير عارفين بوجودها أصلا رغم وجودها الظاهر للعيان على مستوى الواجهات الرئيسية لأغلب المحركات .

الإنترنت والمكتبات.. مقاربات أولية

فإن أخذنا هذا العنصر من الناحية التقنية لربما أرجعناه إلى محور بحثي هام يتعلق بموضوع الواجهات المعلوماتية بين الجهاز والمستفيد (Interfaces Homme-Machine) ولربطناه كذلك أيضا بمجال الدراسات النفسية الحسية (Psychocognitives) التي لها علاقة قصوى بمدى تفاعل المستفيد مع واجهات المحركات. تلك الواجهات التي صارت تحت وطأة المنافسة والإشهار مكتظة بالروابط وغير متناسقة مع القدرات الحسية البصرية للمستفيد الذي عادة ما يركز في هذه الحالات على المناطق الرئيسية للواجهة مثل مناطق طرح السؤال دون السعي إلى التثبت و التمكن من بقية الخدمات البحثية المتوفرة له على الشاشة نفسها. وهناك دراسات عدة في هذا المجال تعنى بدراسة هيكلية الواجهات التطبيقية للبرامج الموجهة للمستفيدين وتبين مراكز تفاعلها مع هؤلاء.

أما إذا أخذنا المسألة من ناحية عملية، فالانعكاسات المباشرة على المستفيد كبيرة إذ إنه يتبين من خلال ما قدمناه أهمية ما يتركه المستفيد جانبا من خدمات ذات قيمة إضافية من الضروري أنها تساعد في تحسين نوعية المعلومات المسترجعة و بالتالي قيمتها الإعلامية و جدواها الاستراتيجية في أخذ القرار. وهنا تكمن معضلة أخرى تمس قطاع التكوين والتدريب التي تشكل عنصرا آخر يجدر الاعتناء به لتطوير قدرات المستفيدين و قابليتهم لمزيد من التفاعل مع النظم المعلوماتية و واجهات التطبيقات العامة والمختصة.

مراجع :

- 1- انظر كتاب محمد فتحي عبد الهادي. التكشيف لأغراض استرجاع المعلومات. جدة: مكتبة العلم 1982 ص 213.
- 2- The Informant تصفح بتاريخ 11/6/2001 على العنوان <http://informant.dartmouth.edu>
- 3- تمت زيارة الموقع في 14,08,2001 <http://www.tracerlock.com>
- 4- تمت زيارة الموقع في 14,08,2001 <http://www.netmind.com>
- 5- تمت زيارة الموقع في 14,08,2001 <http://www.askjeeves.com>
- 6- ولفرد لانكستر. نظم استرجاع المعلومات. ترجمة د. حشمت قاسم. القاهرة: مكتبة غريب، 1981 ص 527 .
- 7- Netscoring santé (تمت زيارة الموقع في 27/08/2001) <http://www.netscoring.com>