أضواء على التطورات المعاصرة
ومستقبل خدمات التكشيف والاستخلاص التقليدية

إعداد

أ.د. أحمد بدر
أستاذ المكتبات والمعلومات
جامعة القاهرة

د. ناريمان إسماعيل متولى
أستاذ علم المعلومات المساعد
كلية الآداب – جامعة الإسكندرية

مقدمة :

يبدو أن بحوث التكشيف والاستخلاص فى تناقص مستمر خلال السنوات الأخيرة ، على الرغم من الحاجة المتزايدة لتحسين نوعية التكشيف والاستخلاص آخذين فى الاعتبار فاعلية التكلفة . ولقد تركزت معظم جهود تطوير البحوث فى المجال ، على كيفية استبدال المصادر المحسبة بالمصادر الإنسانية ، نظراً للفيض الهائل من المعلومات المنتجة كل يوم ، أى أن التركيز هو على استخدام النظم الآلية وليس على التعرف على احتياجات المستفيدين بمستوياتهم المعرفية المختلفة من الكشافات والمستخلصات .

وتتناول هذه الدراسة فى البداية خريطة بحوث التكشيف والاستخلاص شاملة للجوانب المعرفية والتحكم فى المصطلحات والتكشيف والاستخلاص بمعاونة الحاسب وبحوث عرض وتركيب الكشافات ثم تقييم الكشافات والمستخلصات الآلية ثم تتناول الدراسة نماذج من أهم كشافات الوثائق المنتجة آلياً والتى ينتظر لها الذيوع والانتشار فى القرن الحادى والعشرين ثم يقوم الباحثان بتحليل نماذج من الإنتاج الفكرى الحديث (١٩٩٧-٢٠٠٠م) للتعرف على اتجاهاته ولغاته واهتمامه بالتفاعل مع التخصصات الأخرى خصوصاً الحاسبات واللغويات والذكاء الاصطناعى والنظم الخبيرة وبعض الاتجاهات العامة فى النظرية والممارسة وأخيراً النتائج والتوصيات لدعم وتطوير البحوث فى المجال باللغة العربية .

خريطة بحوث التكشيف والاستخلاص :

يركز هذا الجزء على التكشيف باعتباره الخطوة الأولى لإعداد المستخلصات ، وإن كان الباحثان سيتناولان الاستخلاص بالتحديد فى بعض المواقع وعلى غير المتوقع فإن عدد بحوث التكشيف لا تتزايد مع مرور السنين ولكنها لا تقل وقد أكدت هذا الاتجاه الباحثة جيسكا مايلستد (Milstead, J., 1994, p. 577) فقد جاء فى دراستها أن مجلة المراجعة السنوية لعلم وتكنولوجيا المعلومات

obeikandl.com

(ARIST) فى سنواتها العشر الأولى (١٩٦٦-١٩٧٥) قد نشرت فى المتوسط (٢٥-٣٠) مدخل كشفى كل عام تحت رأس موضوع التكشيف والإحالات الموضوعية ، ثم انخفض هذا المتوسط السنوى إلى (١٠-١٢) مدخلاً كل عام ، وإذا كان من العسير وضع التبرير الحقيقى لهذا الانخفاض التدريجى للبحوث ، فهناك من يذهب إلى أن معظم بحوث ودراسات التكشيف والاستخلاص نظرية ، وليس فيها نظم تطبيقات فعلية تستجيب لاحتياجات الحياة اليومية إلا أقل القليل ، أما خريطة بحوث التكشيف المطلوبة من الناحيتين النظرية والعملية فتتضمن ما يلى :

١ - الجوانب المعرفية Cognitive aspects

وهذه الجوانب تتضمن :

(أ) عمليات معرفية فى التكشيف أى التعرف على ما تقوله الوثيقة aboutness والأسئلة التى يمكن أن تجيب عنها ، وترجمة ذلك فى مصطلحات يستخدمها نظام التكشيف سواء كان النص الحر أو المصطلحات المحكومة . والفجوة هنا هو عدم إمكانية الحاسب فى إنتاج كشافات تضاهى ما ينتجه الإنسان . فضلاً عن اختلاف اثنين من الناس عند قيامهم بتكشيف نفس الوثيقة نظراً لاختلافهم فى مستوى فهم الموضوع أو عدم استيعابهم لدلالات المصطلحات أو غير ذلك ، وهذا اختلاف فى الممارسة المعرفية .

(ب) عمليات معرفية فى البحث : أى العمليات التى تخص المستفيدين من الكشاف أى مدى الإفادة Usability ، كيف يقوم الباحثون بصياغة أسئلتهم ، كيف تساعد الكشافات فى تحديد احتياجاتهم الفعلية ؟ إن تنظيم الكشافات لتتلاءم مع العمليات المعرفية للمستفيدين تتناول قضايا الإخراج وتركيب النص Structure وترتيبه وتجميع المفاهيم المتقاربة مع بعضها .

٢ - التحكم فى المصطلحات Vocabulary Control

(أ) علاقات المكنز : هل المستفيدون يعرفون معانى مختلف هذه العلاقات .(BT/NT/RT USE/UF) ... ولا نستطيع القول فى الوقت الحاضر ، رغم كل تطورات استخدام اللغة الطبيعية أننا نستطيع أن نستبعد المصطلحات المحكومة – أى أن الوقت مازال بعيداً قبل أن نجزم بأن محركات بحث استرجاع النص الكامل ستكون من الدقة بحيث تستبعد تماماً التكشيف الإنسانى .

(ب) حجم المكنز : لقد وصلت بعض المكانز إلى حجم ضخم كالمكنز الهندسى Engineering Thesaurus (١٦٠,٠٠٠ مصطلح) فما الحجم المناسب للمكنز فى علاقته بقاعدة البيانات ؟ وهل هناك توازن يمكن الوصول إليه بين استخدام المكنز ودقة الاسترجاع ؟ وعند أى نقطة سيصبح المكنز ضخماً للدرجة التى تكون فيها الوثائق ذات الموضوع الواحد مشتتة تحت مصطلحات مختلفة .

(جـ) التوازن بين التحكم فى المصطلحات واللغة الطبيعية : البحث هنا عند التوازن بالنسبة

للكشاف الالكترونى لتحقيق هذا التوازن ، خصوصاً وهناك من الدراسات التى تشير إلى أن أفضل استرجاع هو مزيج من التكشيف الإنسانى باستخدام المصطلحات المحكومة والقدرة على بحث النص بالكامل ، ومع ذلك فالاتجاه الآن نحو زيادة قدرات برامج الاسترجاع .

٣ - التكشيف والاسترجاع بمعاونة الآلة

Machine aided Indexing and Retrieval

(أ) نظم المصطلحات المقترحة Candidate Term Systems : يستخدم عديد من منتجى قواعد البيانات برامج التكشيف التى تستعين بالحاسب، وهذه البرامج تقوم بتحليل النص (عادة المستخلصات) ثم تقترح مصطلحات ممثلة لموضوع النص ، وهنا تقع المسئولية على المكشف لمراجعة هذه المصطلحات المقترحة ، يحذف منها أو يضيف إليها ما يراه ، ومن أمثلة هذه النظم المطبقة فعلاً ماهو موجود فعلاً فى معهد البترول الأمريكى ، وفى مركز المعلومات الفنى الوطنى ، وفى ناسا (NASA) .. وهناك بعض الشركات الخاصة التى قامت مؤخراً بتقديم خدماتها فى هذا التكشيف الذى تعاونه الآلة .. ولكن ما أثر هذا الاتجاه على نوعية التكشيف ، وإلى أى مدى يقترح النظام المصطلحات المفيدة التى ربما فاتت على المكشف اليدوى ؟ وما تكاليف هذه النظم على المدى البعيد ؟

(ب) استكمال التكشيف الإنسانى Supplementing Human Indexing : هناك اتجاه آخر لا يبدأ من الآلة ، بل من الإنسان حيث يقوم المكشف بإعداد مدخلاته من المصطلحات ، أى يتم اقتراح مصطلحات عليها عن طريق الحاسب ، وتعتبر المكتبة الوطنية الطبية واحدة من أهم المؤسسات التى تطبق هذا النظام والمسمى (Med In Ex) والمهم أن فاعلية التكلفة الخاصة بهذه النظم غير معروفة تماماً (Milstead, J., 1994, p. 580) .

(جـ) الجمع بين التكشيف اليدوى واسترجاع اللغة الطبيعية : هناك نظم فى الوقت الحاضر تجمع بين التكشيف الإنسانى والأساليب الإحصائية والدلالية Semantic والنحوية Syntactic للاسترجاع ، آخذة فى اعتبارها فاعلية التكلفة ، وواضح أن هذه النظم تعمل اعتماداً على بعض الجهد الإنسانى فى التحليل، ولكن ما الأجزاء التى يمكن أن يحل محلها استرجاع اللغة الطبيعية حتى يمكن الوصول إلى نتائج تعاونية ناجحة ؟ مازال الأمر قيد الدراسة والبحث .

(د) نظم الفئات Categorization Systems : هناك دراسات جارية تتعلق بتنظيم النصوص إلى مجموعة فئات Categories محددة سلفاً ، ويذهب القائمون على هذه النظم بإمكانية الوثوق بها (مثل المتبع فى النظام الدولى للمعلومات الذرية) وهذا النظام هو جزء من عملية ضبط الجودة حيث يستخدم مجموعة من القواعد لضبط الفئات مع مصطلحات التكشيف ، وإذا ما حدث خلل فى تشغيل النظام فيرجع إلى المكشف الإنسان .

٤ - العرض والتركيب

Layout and Structure

تنسحب المناقشة هنا على الكشافات المطبوعة والإلكترونية . كما أن الأمثلة التى ستعرض هنا هى مجرد نماذج وليست حصراً للمجال .

(أ) الإحالات Cross-references : هل يدرك المستفيدون حقاً المقصود بإحالة «انظر» «انظر أيضاً» والفرق بينهما وما أقصى عدد من هذه الاحالات يكون مفيداً ، أى الحد الذى سيقف عنده الباحث عن الاستعانة بهذه الأدوات ؟ خصوصاً وقد أظهرت بعض البحوث الارتباك والإحباط الذى أصاب بعض المستفيدين من جراء هذه الإحالات ، وفى قاعدة البيانات ما قيمة الإحالات وكيفية القيام بها بكفاءة ؟ وفى كشافات الكتب هل يستخدم التكشيف انظر أيضاً التكشيف الآلى: التكشيف المتسلسل ،
أو التكشيف
الآلى انظر التكشيف الآلى ،
المسلسل انظر التكشيف المتسلسل .
وبالمثل إحالة انظر أيضاً :
التكشيف ٣٢ ، ٥٦ انظر أيضاً «التكشيف الآلى» ،
البحث ١٨٣-١٨٥ ،
التدريب ٧٤،
أو
التكشيف ٢٣ ، ٥٦ ،
البحث ١٨٣ - ١٨٥ ،
التدريب ٧٤ ،
انظر أيضاً التكشيف الآلى .

(ب) الترتيب : arrngement هل يكون الترتيب كلمة بكلمة أو حرف بحرف ، ومع هذين الاختيارين ما نوع المدخل (بالمكان أو بالموضوع) أو يكون الترتيب على أساس الحروف فقط (ليس ما يسمى بالدليل الأمبيريقى، وهناك حالات عديدة تتعلق بالمدخل (المكون من عدة كلمات بعضها محول أو مبين (مثل كلمة بين Between , among ، تطبيق ، من/فى in/from) فهل يتم تجاهل الحروف والاكتفاء بالأسماء ؟

(جـ) الطول Length : إلى أى مدى يكن طول الكشافات أو المستخلصات بالنسبة لنص الأصلى ؟

٥ - تقييم التكشيف والاستخلاص

(أ) تقييم الاسترجاع :

الطرق المتاحة لتقييم نوعية الاستخلاص تتركز فى منهجية الاستدعاء والدقة Recall and Precision وتفريعاتها ، وعلى الرغم من مزايا هذه المنهجية ، إلا أن لها عيوبها الواضحة ، ولعلنا نذكر منها اثنين هما :

- قياس الاستدعاء فى المجموعات الحقيقية تتطلب تقييم كل وثيقة فى المجموعة ومدى صلاحيتها relevance لكل سؤال ، وهى مهمة مستحيلة .. والدراسات التى تدور حول المشكلة بالبحث من عدة وجهات نظر ، أو البحث بطريقة عريضة Broad ثم تجميع كل الاسترجاعات ذات الصلة relevant لتزويدها بقاعدة الوثائق الصالحة .. هذه الدراسات ليس لديها أى وسيلة لتخمين الوثائق التى يمكن أن تكون ناقصة .

- قياس الصلاحية relevance measure هو قياس ناقص فى ذاته ، فمن الناحية العملية يعادل المقياس كشيء مطلق ، ويعتمد هنا على درجة المضاهاة بين ماذا تتحدث عنه الوثيقة Aboutrness ومصطلحات السؤال ولكن هناك عوامل تدخل وتؤثر على فائدة وثيقة معينة فى موقف معين . فعلى سبيل المثال لا ينبغى ألا تضاهى الوثيقة مستوى معرفة المستفيد ، ولكن قيمتها تعتمد جزئياً على ما يعرفه المستفيد مسبقاً ، فإذا كانت هناك وثيقتان متشابهتان ، وقد تكون الأولى بالنسبة للمستفيد مفيدة جداً، فقد تكون الثانية ذات أهمية متدنية لأنها لا تضيف إلا القليل لمعرفة المستفيد . هناك حاجة بحثية ماسة إلى مقاييس أخرى فى المواقع العملية لتحقيق نتائج موثوق بها .

(ب) اختبار مدى استخدام Usability Testing : اختبار الاستخدام عملية مكلفة للغاية ، وهامش الربح عن مبيعات الكشافات ليس كافياً لدعم هذا الاختبار ، ومع ذلك فالبحث فى هذا المجال مفيد للغاية .

(جـ) قيمة التكشيف : القضية التى تتعلق بها مختلف الأطراف هى الرد على سؤال مهم هو : هل يستحق التكشيف كل هذا الجهد ؟ كيف يمكن قياس القيمة فعلاً ؟ ، وإذا كان التكشيف الإنسانى مكلف فالضغوط مستمرة لتقليل هذه التكاليف .. ولكن ما القيمة الفعلية للكشاف ، هل كشاف الكتاب يزيد من قيمته ؟ وإذا كانت الزيادة صحيحة فكيف يمكن قياسها ، والسؤال نفسه ينطبق على النص الكامل والمستخلصات ، والأهم هنا أيضاً هو حساب القيمة المضافة added value فى حالة وجود الفهرس . وإذا كان التقييم يتصل بجوهر مشكلة الاسترجاع فهى قضية فى غاية الصعوبة فى الدراسات البحثية .

انتاج وتوليد كشافات الوثائق آلياً على مشارف القرن الحادى والعشرين

(أ) تقديم :

لقد أدت الزيادة الهائلة فى كمية المعلومات المتاحة من خلال الوسائط الإلكترونية (بما فيها الإنترنت ، إلى زيادة الأهتمام بإنشاء وتطوير طرق آلية للقيام بعملية الاقتباس Extracting أو الترشيح filtering أو التلخيص summarization أو استرجاع المعلومات ذات العلاقة بمهام محددة ومعظم هذه المعلومات موجودة فى شكل وثائق إلكترونية كالمقالات والتقارير (انظر Belkin; Conie; Cardie, Salton) ويلاحظ القارئ استخدام مصطلحات الاقتباس والترشيح والتلخيص فضلاً عن الأساليب الفنية التى تمت بالنسبة لاكتشاف المعرفة Knowledge discovery فى قواعد البيانات التى تستخدم أيضاً عملية اقتباس المعلومات من البيانات النصية . أى أن كلمة مستخلصات واستخلاص لم تعد تستخدم إلا نادراً مع النظم الالية .

وتشير عملية اقتباس المعلومات إلى اقتباس ملخص المعلومات من النص بحيث تصفه فى شكل تركيبى Structured Form (Cardie, c 1997, p. 65) أما عملية ترشيح المعلومات واسترجاعها فتهتم بتحديد البيانات غير التركيبية أو نصف التركيبية .

والتركيز فى عملية ترشيح المعلومات هو على اهتمامات المستفيد user profiles أما اكتشاف المعرفة فهو مصطلح حديث نسبياً يستخدم مع قواعد البيانات التركيبية (وعادة العلاقية Relational) ويمكن لأساليب اكتشاف المعرفة أن تستخدم مع البيانات النصية أيضاً فضلاً عن استخدام أساليب اكتشاف المعرفة لتصنيف النص إلى موضوعات Topics (حسب تصنيف مكتبة الكونجرس مثلاً..) .

(ب) أنواع نظم التكشيف الآلية الحديثة وكيفية تقييمها:

قامت الباحثة جوليا هودج (Hodges, J., 2000) باستعراض لأهم مداخل نظم التكشيف الآلية من النواحى الفكرية ، ثم أوردت نماذج منها وكيفية تقييمها ، ويلاحظ هنا أن تقييمها اعتمد على طرق الاستدعاء والدقة (recall & precision) على الرغم مما سبقت الإشارة إليه فى خريطة التكشيف إلى سلبيات هذه الطرق التقييمية وهى تذكر أن هناك نظم تكشيف آلية تستخدم النماذج الاحتمالية probabliatie models وهى التى تقوم بتقدير الوزن الاحتمالى لمختلف الكشافات باستخدام عدد من الأسئلة المتعلقة بوثيقة محددة كإحدى الطرق اللازمة للتعرف على الصلاحية Relevance.

وهناك نظم أخرى تجمع بين التوزيع الإحصائى والتحليل اللغوى Linguistic heuristics لتحديد الكشافات الصالحة للوثيقة ، وهناك مداخل أخرى تفيد من تجهيز اللغة الطبيعية (NLP) بالتعرف على جذور الكلمات Stemming (أحياناً تسمى الجذر root والجذع stem) كما تحاول نظم أخرى الإفادة من معلومات المتن (أو المضمون) Contextual information لتكشيف الوثائق .

وترى جوليا هودج أن معظم التركيز الحالى فى بحوث التكشيف الآلى هو تكشيف مفاهيم conceptual indexing (أى ما يسمى أيضاً بالتكشيف بالتعيين assignment indexing) فى مقابل نظم التكشيف المعتمدة على الكلمات word indexing من النص عن طريق الاقتباس Extraction وهى تؤكد ضرورة أن تكون نظم التكشيف ممثلة لدلالات semantics محتوى المقال وليس مجرد الطرق التركيبية Syntactical ways ذلك لأن التكشيف المفهومى يساعد عملية استرجاع المعلومات لسببين على الأقل أولهما أن هناك مفاهيم أقل من إمكانيات التمثيل التركيبى rsyntactic representation وثانيهما أن الكشافات المفهومية يمكن أن تقدم لنا المعلومات الدلالية semantic المهمة وذلك عندما يحاول نظام استرجاع المعلومات لمضاهاة السؤال بالمعلومات فى الوثيقة . ولكن التكشيف المفهومى يحتاج إلى المعرفة الواسعة بالتخصص العلمى فى المجال .

وقد قامت جوليا هودج باستعراض نماذج من نظم التكشيف المختلفة ومداخلها البحثية ، وهى :

١ - نظام سمارت SMART
(Salton Magic Automatic Retriever of Text)

٢ - نظام ايمز AIMS
(Assisted Indexing Mississippi State)

٣ - نظام ديدال Dedal وهو كشاف مفهومى للنص والرسومات ووثائق التسجيلات videotaped فى حقل الهندسة الميكانيكية .

٤ – نظام وورلدفيو Word View وهو يقوم بتجهيز ومعالجة مقالات الأخبار والتقارير الفنية من المؤسسات المختلفة مثل معامل بيل Ball Lab .

٥ – نظام سيزور Scisor System (System for conceptual Information Summarization, Organization and Retrieval)

وهو يقوم بتحليل القصص الإخبارية خصوصاً الأخبار المالية .

٦ – نظام كوسكوت Quescot System (Quantification and encapsulation of semantic context)

واعتمدت فى التقييم على مقاييس الاستدعاء والدقة حيث :

$$\text{الاستدعاء} = \frac{\text{عدد الكشافات الصحيحة التى تم توليدها}}{\text{العدد الكلى للكشافات الصحيحة}}$$

$$\text{الدقـــة} = \frac{\text{عدد الكشافات الصحيحة التى تم توليدها}}{\text{العدد الكلى للكشافات التى تم توليدها}}$$

وقد كان أفضل معدل استدعاء ودقة تم الوصول إليها عن طريق خبراء ما يسمى بمؤتمر التفاهم للرسالات (MUC) هو حوالى ٤٠ ٪ – ٥٠ ٪ (Cowie, J., 1996) وهذه نظم تقوم بالاقتباس والتلخيص ، ولكن نظام ايمز AIMS كان له معدل دقة يصل إلى (٦٢ – ٦٦,٥ ٪) ومعدل استدعاء (٧٢ – ٩٠ ٪) (Hodges, J., 2000, p. 15) .

وقد استخدم النظام عدة تجارب تفيد من مختلف أجزاء المقال :

- المستخلص وحده .
- المستخلص والمقدمة .
- المستخلص وجزء من المقدمة .
- المستخلص والخاتمة Conclusion .
- المستخلص وجزء من المقدمة وجزء من الخاتمة .

ثالثا : تحليل لبعض التطورات المعاصرة للتكشيف والاستخلاص فى الانتاج الفكرى :

اختار الباحثان عشرين مقالاً من مجلة الاتجاهات المكتبية Library Trends (1999) ومن موسوعة المكتبات والمعلومات ELIS عام (٢٠٠٠م) ومن مجلة الجمعية الأمريكية لعلم المعلومات JASIS ، ومعروف أن الأخيرة أهم الدوريات فى مجالها على المستوى العالمى . ولعل استعراض المقالات السابقة والمنشورة ما بين الأعوام ١٩٩٧/٢٠٠٠م يظهر لنا أضواء وجوانب بعض التطورات المعاصرة والمستقبلية فى التكشيف والاستخلاص .

نماذج من البحوث والدراسات بمجلة الجمعية الأمريكية لعلم المعلومات (JASIS) :

1. Disciplinary Variation in Automatic Sublanguage Term Identification.

(١) يصف البحث طريقة لتحديد مصطلحات المجال الموضوعى الأساسية والموجودة باللغة البديلة (SL) وبالطريقة الآلية ، وتعرف اللغة البديلة (SL) بأنها اللغة المستعملة فى مجال موضوعى محدد بواسطة المتخصصين .

- مدرسة علم المعلومات والمكتبات ، جامعة شمال كارولينا (شابيل) .

2. A Statistical Learning Approach to Automatic Indexing of Controlled Index Terms.

(٢) مدخل إحصائى تعليمى للتكشيف الآلى الخاص بالمصطلحات المحكومة للكشافات ويهدف البحث تحسين أداء عملية التعلم ، وهذه بدورها ستؤدى إلى تحسين أداء التكشيف .

- قسم علم الحاسب الآلى ، الجامعة الصينية فى هونج كونج .

3. A Concept Space Approach to Addressing the Vocabulary Problem in Scientific Information Retrieval : An Experiment on the Worm Community System.

(٣) استخدام مدخل سعة المفهوم لحل مشكلة المصطلحات فى استرجاع المعلومات العلمية تجربة على نظام مجتمع الديدان مع التعرف على استخدامات هذا المدخل فى شبكة الإنترنت والمكتبات الرقمية .

- قسم نظم المعلومات الإدارية (MIS) مدرسة الإدارة جامعة اريزونا .
- المدرسة العليا لعلم المكتبات والمعلومات جامعة الينوى (أريزونا – شامبين) .

4. Why Words and Co-Words Cannot Map the Development of the Sciences.

(٤) لماذا لا تؤدى الكلمات والكلمات المتزامنة إلى وضع خريطة لتطور العلوم .

قد أثبت الباحث رسالته هذه باستخدام مقالات النص الكامل فى مجال الكيمياء الحيوية ، وذلك لتحليل الكلمات ومقارنتها بالنسبة لتواجدها أو غيابها المتزامن فى هذه النصوص

- قسم ديناميكية العلوم والتقنية جامعة أمستردام هولندا .

5. Citation Context versus the Frequency Counts of Citation Histories.

(٥) سياق الاستشهاد مقابل حساب التكرار لتواريخ الاستشهاد ، وقد استخدم الباحثون أكثر من مائتى ورقة بحث لتحليلها .. وتبين للباحثين أنه ليس هناك تطابق بين نتائج رتبة سياق الاستشهاد عند مقارنتها برتبة موضوعات البحوث عن طريق حساب تكرار استشهاداتها .

- وحدة دراسات العلوم – كلية الطب – جامعة زغرب ، كرواتيا .

6. Professional Summarizing : No Cognitive Simulation without Observation.

(٦) التلخيص المهنى : ليس هناك محاكاة معرفية دون الملاحظة .. يشير الباحث أن التلخيص المهنى يتضمن عمليات معرفية كالاستخلاص والتكشيف والتصنيف ، والتى يقوم بها الخبير لإعداد نظم المعلومات الببليوجرافية والتى تمثل صيغاً مصغرة للقارئ وتعتبر مرشحات للمعلومات، ولكن العمليات المعرفية تتم داخل رؤوسنا ، وبالتالى فهى تستبعد الملاحظة المباشرة وينادى الباحث بضرورة دراسة التجهيزات الفنية والمنهجية اللازمة للإفادة من الملاحظة .

- قسم المعلومات والاتصال معهد البوليتكنيك – هانوفر – المانيا .

7. Khan, Kushal and Craig Locatis (1998, 172-182) Searching through Cyberspace : The Effects of Link Display and Link Density on Information Retrieval from Hypertext on the World Wide Web (JASIS, 49(2)).

(٧) البحث خلال الفراغ الكونى : تأثير روابط العرض والكثافة على استرجاع المعلومات من الهيبرتكست على الشبكة العنكبوتية العالمية .. تتناول الدراسة استرجاع المعلومات من الهيبرتكست على الشبكة العنكبوتية العالمية للإنترنت ، وقد تبين وجود تأثيرات ذات دلالة لكل من كثافة الروابط (عدد الروابط لكل عرض) وشكل العرض (فى الفترات والقوائم) وذلك بالنسبة لنتائج الأداء فى البحث ، كما يعتمد الأداء على الخبرة المسبقة فى تصفح واستعراض الهيبرتكست على الشبكة العنكبوتية .

- قسم العلم المعرفى (Cognitlive Science) المكتبة الوطنية الطبية – أمريكا .

8. Khan, Kushal (1998). Searching through Cyberspace: The effects of Link Cues and Correspondence on information Retrieval from Hypertext on the world Wide Web. (JASIS, 49 (14) 1248-1253).

(٨) البحث خلال الفراغ الكونى : تأثير الروابط المعبرة ، وروابط المراسلة على استرجاع المعلومات من الهيبرتكست على الشبكة العنكبوتية العالمية .. تتناول الدراسة تأثير الروابط المعبرة وروابط مراسلة الكلمات على الأداء البحثى باستخدام وثيقة هيبرتكست محورية ولها روابط خارجية بمصادر أخرى على الشبكة العنكبوتية العالمية .

- جامعة إلينوى والمكتبة الوطنية الطبية .

9. Clumping Properties of Content - Bearing Words.

(٩) الصفات التجميعية للكلمات الحاملة للمحتوى الموضوعى المميز .. تعتمد هذه الدراسة على ما تقوم به نظم استرجاع المعلومات من التعرف على الكلمات التى تحمل محتوى موضوعيًا مميزاً ، كما قد تعين هذه النظم أوزاناً لهذه الكلمات كجزء من عملية صياغة الأسئلة . وتحدد هذه الدراسة فكرة التجمع Clustering المتسلسل للكلمات فى النص كمؤشر للمحتوى الذى تحمله الكلمات . وقد تم تجميع الأدلة التجريبية من قواعد معلومات النص الطبيعى ومن ثلاث لغات مختلفة .

- مركز دراسات المعلومات واللغات ، جامعة شيكاغو – أمريكا .
- قسم الرياضيات ، جامعة بارلان Barllan – إسرائيل .
- قسم علم الحاسب ، جامعة توركو Turku – فنلندا .

10. Price Index and Its Relation to the Mean and Media Reference Age.

(١٠) كشافات برايس وعلاقته بمتوسط ووسيط العمر المرجعى .. يتضمن المقال فى جزئه الأول اثبات أن كشاف برايس يعتبر دالة لمتوسط ووسيط العمر المرجعى ، أما الجزء الثانى فيشير إلى العلاقة المحددة والتى يتم شرحها بواسطة نظرية الاحتمالات والاحصاء .

- جامعة ليمبرج Limburg – بلجيكا .

11. Interaction in Information Retrieval : Selection and Effectiveness of Search Terms.

(١١) الروابط البينية فى استرجاع المعلومات : اختيار وفاعلية مصطلحات البحث .. تناولت الدراسة مصادر وفاعلية مصطلحات البحث المستخدمة أثناء البحث على الخط المباشر وبحضور وسيط معلوماتى ، وكان عدد المستفيدين (أربعين) وعدد الوسطاء (أربعة) حيث كان المستفيد يقدم سؤالاً ويتفاعل مع الوسطاء الأربعة ، وذلك باستخدام نظام ديالوج . وكانت أكثر المصادر فاعلية تلك التى تعتمد على مصطلحات البحث المكتوبة فى أسئلة الباحثين فضلاً عن المصطلحات المتعلقة فى التغذية المرتدة .

- قسم علوم المكتبات والمعلومات ، جامعة شمال تكساس – أمريكا .
- مدرسة دراسات الاتصال والمعلومات والمكتبات، جامعة رانجرز .

12. Design and Implementation of Automatic Indexing for Information Retrieval with Arabic Documents.

(١٢) تصميم وتطبيق التكشيف الآلى لأغراض استرجاع المعلومات باستخدام الوثائق العربية.. استخدمت فى الدراسة عدد (٢٤٢) مستخلصاً للوثائق العربية فى مجال نظم [illegible]ت وعلم الحاسب الآلى ، كما تم تصميم وبناء نظام استرجاع معلومات آلى لمعالجة البيانات العربية ، واستخدام النظام لغة (C) والمرجم (GCC) على IBM/PC حاسبات مصغر متلائم ، وأثبتت التجربة أن التكشيف الآلى له نفس فاعلية التكشيف اليدوى .

- معهد الأردن للعلوم والتقنية – أربد – الأردن.
- قسم الحاسب الآلى – معهد إلينوى للتقنية – شيكاغو .

13. Experiments with Automatic Indexing and a Relational Thesaurus in a Chinese Information Retrieval System.

(١٣) تجارب فى التكشيف الآلى والمكنز العلاقى فى نظام استرجاع المعلومات الصينى ..

تجارب مع نظام استرجاع تفاعلى (CIRS) ومكنز علاقى تفاعلى .. تقاس الفاعلية بواسطة الاستدعاء والدقة ، وقد تبين بالتحليل الإحصائى أن قياسات الاستدعاء ، الدقة تتحسن بواسطة المكنز العلاقى سواء كان ذلك فى التكشيف الآلى أو اليدوى .

- قسم علم الحاسب الآلى – معهد إلينوى للتقنية – شيكاغو .

14. Indexing and Access for Digital Libraries and the Internet : Human, Database, and Domain Factors.

(١٤) التكشيف والإتاحة للمكتبات الرقمية والإنترنت : العوامل الإنسانية وقواعد المعلومات والمجال العلمى ..

دراسة لبعض العوامل والقضايا التى تؤثر على التكشيف والإتاحة عند تصميم النظم لاسترجاع المعلومات فى المكتبات الإلكترونية الرقمية المستقبلية والإنترنت . فالمستفيدون يبحثون فى كيفية الوصول إلى أفضل النتائج على أسئلتهم عند البحث على الشبكة العنكبوتية العالمية خصوصاً وهناك من الباحثين القريبين لعلم المعلومات (كالعلم اللغوى واللغويات المحسبة والذكاء الاصطناعى) ممن يدرسون قضايا استرجاع المعلومات من جهة نظرهم ، وتشمل العوامل الإنسانية : البحث الموضوعى بالمقارنة بالتكشيف والمصطلحات المتعددة للإتاحة ، وكذلك التصنيف الذى يتبعه الشخص العادى Folk Classification أما عوامل قاعدة البيانات فتشمل قانون برادفورد والمصطلحات المساعدة Vocabulary Scalability وقاعدة ريزينكوف - Resnikoff Dolly 30 : 1 Rule أما عوامل المجال العلمى فتشمل دور المجال فى التكشيف .

• قسم دراسات المعلومات – جامعة كاليفورنيا – لوس أنجلوس .

15. He, Qin (Summer 1999).

مجلة الاتجاهات المكتبية (عام ١٩٩٩)

Knowledge discovery through Co - word Analysis

library Trends. V 48 (1), p. 133-159

(١٥) اكتشاف المعرفة من خلال تحليل لمصاحبة الكلمات .. وتتناول الدرسة تحليل مصاحبة الكلمات كأسلوب للتعرف على عدد مرات الحدوث المشترك Co-occurance لزوج من الكلمات أو الفقرات ، وذلك لاكتشاف الروابط بين موضوعات البحوث ، وبالتالى تتبع تطور العلم فى مجال معين ، وذلك يؤدى إلى التخطيط الأفضل للبحوث العلمية ، الأمر الذى يجعل هذا الأسلوب أداة قوية لاكتشاف المعرفة فى قواعد البيانات .

16. Qin, Jian (Summer 1999), Discovering Semantic Patterns in Bibliographically Coupled Documents. Library Trends, V. 48 (1), p. 109-132.

(١٦) اكتشاف النماذج الدلالية فى الوثائق المتزاوجة ببليوجرافيا .. فتتناول الدراسة قضايا اكتشاف المعرفة فى قاعد البيانات الببليوجرافية ، وقد استخدم مثال لتحليل النموذج الدلالى للتعريف بالجوانب المنهجية لاكتشاف المعرفة فى قواعد البيانات اعتماداً على الكلمات المفتاحية المختارة من الوثائق والمجمعة بواسطة التزاوج الببليوجرافى .

17. Ahonen, Helena (Summer 1999), Knowledge Discovery in Documents by Extracting Frequent Word Sequences. Library Trends, V. 48 (1), p. 160-181.

(١٧) اكتشاف المعرفة فى الوثائق عن طريق اقتباس عدد مرات تتابع الكلمات . تتناول الدراسة طريقة لاقتباس أكثر الكلمات المتتابعة تردداً فى الوثائق حيث تحاول طرق اكتشاف المعرفة الكشف عن النماذج والانتظامات العامة فى البيانات بدلاً من الكشف عن حقائق معينة .. وهذه الطريقة عسيرة للتطبيق بالنسبة للإنسان ولكنها ممكنة آلياً .

18. Pinto, Maria and Lancaster, F.W. (Summer 1999). Abstracts and Abstracting in Knowledge Discovery. Library Trends V. 48 (1), pp. 234-248.

(١٨) المستخلصات والاستخلاص وأهميته فى اكتشاف المعرفة .. تتناول هذه الدراسة متطلبات المستخلصات التى يمكن قراءتها بواسطة الإنسان ومقارنتها بالمتطلبات الخاصة بالبحث بالحاسب الآلى ، وينتهى الباحثان إلى أن الإتاحة الواسعة للنص الكامل الإلكترونى ، لا تقلل من قيمة أنشطة المستخلصات لاسترجاع المعلومات حتى فى التطبيقات الأكثر تعقيداً فيما يعرف الآن باكتشاف المعرفة Knowledge Discovery .

19. Chowdhuny, Golinde G. (Summer 1999), Template Mining for Information Extraction from Digital Documents, Library Trends, V. 48 (1), p. 182-208.

(١٩) استخراج اقتباسات المعلومات من الوثائق الرقمية .. تتناول الدراسة عملية اقتباس المعلومات آلياً من نصوص اللغة الطبيعية اعتماداً على مدخل الوسادة فى Template ، وذلك عند تشكل البيانات أو النص المحيط بها نماذج يمكن التعرف عليها ، ويستخدم فى هذه العملية محركات بحث الوب Search Engines مثل التافيزتا وذلك لمعاونة المستفيد النهائى فى توليد الاقتباسات من الوثائق الرقمية .

- نموذج من بحوث عام (٢٠٠٠) فى موسوعة علم المكتبات والمعلومات .

20. Hodges, Julia (2000).

مقالة الموسوعة

Automated System for the Generation of Document Indexes. In : Encylopedia of Library and Information Science, V . 66, Suppl. 29, pp. 1-19.

(٢٠) نظام آلى لتوليد كشافات الوثائق ..

يتناول هذا المقال الاستعراضى المطول إنتاج الكشافات بنظام آلى ، ويركز على أن البحوث المعاصرة فى التكشيف الآلى هى بحوث فى التكشيف المفهومى Conceptual Indexing بالمقابلة بنظم التكشيف التى تعتمد على الطرق النحوية (التركيبية أو الصرفية) إلى الاعتماد على الكلمات فى النص ، أى أن التركيز الحالى هو على الكشافات التى تمثل دلالات Semantics المحتوى فى المقال ، وليس على الطرق التى يتم بها تمثيل تلك المفاهيم نحوياً فى النص .. ويركز مقال الموسوعة أعلاه على كيفية Nانتاج الكشافات الآلية للوثائق فى شكل مقروء آلياً وباللغة الطبيعية .

رابعاً: اتجاهات في النظرية والممارسة :

١ - من الإنتاج الفكرى المطبوع إلى الإنترنت :

تناول الباحث وايلى التطورات التى حدثت بالنسبة لخدمات التكشيف والاستخلاص وتساءل هل هناك مستقبل لهذه الخدمات التقليدية المطبوعة؟ وفى إجابته عن هذا التساؤل ذهب وايلى إلى أن عديداً من نظم المعلومات تواجه اليوم خطر التقادم نظراً للتطورات السريعة فى تكنولوجيا الحاسبات والاتصالات ، وأن بعض المنتجين لقواعد البيانات العلمية الرئيسية يقومون فى الوقت الحاضر بتطوير نظم إنتاجية جديدة للاستجابة للاحتياجات المتغيرة للسوق ولمتطلبات الطرق السريعة للمعلومات .

وإذا كان الماضى القريب قد شهد أندماج مختلف التكنولوجيات فى نهاية الستينيات وأوائل السبعينيات بما فى ذلك نظم اختزان البيانات المنخفضة التكاليف ونظم الاتصالات الرخيصة ، فضلاً عن تحسيب العديد من عمليات النشر .. التى أدت بدورها لتطوير نظم البحث على الخط المباشر .. فإن هناك اندماجاً مماثلاً للجيل التالى من التكنولوجيات المتمثل فى الحاسب المنخفض التكاليف ذى القوة الاختزانية الكبيرة ، فضلاً عن نظم الاتصالات السريعة للغاية المسماة بالعميل / الخادم Client / Server وكذلك النشر الإلكترونى على المكتب ، وهذه أدت بنا إلى الطرق السريعة للمعلومات .

ويشهد الوقت الحاضر عديداً من الهيئات التى استكملت إنشاء «النظم الإنتاجية الجديدة» وحولت الهيئات محور اهتمامها من نظام إنتاج للمطبوعات إلى إنتاج مواز للأشرطة اللازمة للبحث على الخط المباشر ، ثم إلى إنتاج ملفات الخط المباشر والتى يمكن أن تنتج بدورها أشكالاً من المطبوعات والمنتجات الإلكترونية .. وهذا مفهوم بسيط ، ولكنه ضخم التعقيد عند التنفيذ . وقد لوحظ أن الشكل الالكترونى للبيانات قد أدى إلى دخل أكبر وأكبر .

ولكن ماذا يحمل المستقبل ؟ الناشرون يرون طلبات أكبر على بياناتهم فى الشكل الإلكترونى وأنهم فى حاجة إلى مرونة لتوصيل المنتجات التى يحتاجها المستهلك أمام المنافسة المتزايدة فى إنتاج المعلومات الالكترونية ولذلك كان الاتجاه نحو الإنترنت (Wiley, D.L., 1994) .

٢ - النظم الانتاجية الجديدة لخدمات التكشيف والاستخلاص :

تنشئ خدمات الاستخلاص والتكشيف نظماً إنتاجية جديدة طبقاً لخطوات منهجية عديدة تبدأ بتشكيل لجان استشارية للتعرف على مختلف جوانب النظام المطلوب من التحرير إلى التسويق إلى التحسيب إلى الإدارة، وعلى سبيل المثال لا الحصر ، فقد قامت هيئة المستخلصات الكيميائية الدولية CABI بتعميم نظامها الإنتاجى الجديد ، وكان عليها أن تتعامل مع تاريخ طويل العدد (٢٦) دورية مستخلصات مستقلة ، وقد كان لديها قاعدتان للبيانات يتم إنتاجهما فى مكانين مختلفين وتستخدمان ممارسات تكشيف مختلفة ، ثم قامت CABI بتحويلهما إلى مدخلات لقاعدة بيانات مركزية للمستخلصات الكيميائية ، والتى يمكن استخلاص مختلف المنتجات منها .

لقد شكلت هيئة CABI لجنة ممثلة لمختلف الأنشطة المتصلة بالتحرير وضبط الجودة والنظم والتسويق والتزويد من أجل تصميم النظام الجديد .. ومع ذلك ومع وجود أكثر من مائة اختصاصى معلومات يقومون بالعمل فى إنتاج قاعدة البيانات فقد كان ذلك تحدياً فى تصميم النظام الذى يرضى الجميع .. وتبنى الجميع مفهوم «التكشيف من أجل قاعدة البيانات» وليس التكشيف لدورية متخصصة بعينها .. وقد أدى ذلك إلى تكشيف أكثر انتظاماً بالنسبة للباحثين للقاعدة على الخط المباشر ، وأدى ذلك التنظيم إلى سهولة إنتاج منتجات جديدة للمستهلكين .

لقد قامت هيئة المستخلصات الكيميائية بإعادة تكشيف قاعدة بيانات المستخلصات الكيميائية والتى تحتوى على أكثر من ٢,٨ مليون تسجيلة .. وقد تم التعبير الفعلى بواسطة برامج الحاسب الآلى ، ولكن سبقها جهد ضخم يتصل بتقرير وتحديد ماذا يتم تغييره وكيفية هذا التغيير ؟

فبالإضافة إلى تصحيح البيانات القديمة ، أضيفت حقول جديدة إلى البيانات الداخلة فى النظام ، وذلك لضمان جودة نوعية المستخلصات ، وتوفير نقاط إتاحة إضافية للمستفيد . ومن الحقول الجديدة التى أضيفت هناك العنوان البريدى للمؤلف Author's e-mail إذا كان فى المقال الأصلى ، وكذلك بعض مصطلحات التكشيف الإضافية للاسترجاع الأكثر دقة .. كما قامت CABI بإضافة حقول جديدة شاملة لحقل الواصفات حتى تمكن المستفيدين من البحث بالمصطلحات العريضة فى مكنز المستخلصات الكيميائية بطريقة هرمية ، ثم الاسترجاع الآلى للمصطلحات الأضيق دون الحاجة إلى اختيارها بطريقة فردية ، كما أضافت فوق هذا كله أكواد تصنيفية جديدة لكل تسجيلة فى قاعدة البيانات .

٣ - التغييرات المعاصرة والمستقبلية والمستفيد :

لقد كانت عين المسئولين وكيفية توصيل هذه التغييرات إليه .. خصوصاً وأنه فى عصر فيضان المعلومات الغامرة Information Overload ، وبالتالى فإن معظم المستفيدين لا يعرفون كيفية بناء قاعدة البيانات ولاكيفية استرجاع المعلومات بالطريقة السليمة فهم يطبعون الكلمات على النهاية الطرفية ثم يتقبلون أى مخرجات يحصلون عليها. ومع ذلك فينبغى التأكيد على أننا فى عصر المعلومات نستطيع أن نبنى أدوات استرجاع بالغة التعقيد وجعلها سهلة الاستخدام بالنسبة للمستفيد، وعلى كل حال فيجب أن يكون المستفيد قادراً على طباعة كلمة أو مصطلح فى تخصصه ، ثم يقوم النظام بطريقة آلية بالبحث عن المصطلحات القريبة أو الشبيهة ، مع تغييرات تتصل بكيفية الهجاء أو كونها جمعاً أو مفرداً .. الخ دون إسهام منه فى ذلك .

ونظم الاسترجاع هذه يمكن أن تتطور فقط مع التكشيف الدقيق والتيجان التى تتوفر فى قاعدة البيانات كإحدى مكوناتها .. وماهو سهل بالنسبة للمستفيد فى الاستخدام يعتبر عادة معقداً للغاية بالنسبة لأولئك الذين قاموا بالإعداد وإنتاج قاعدة البيانات .

٤ - إدخال البيانات مفتاح إنتاج قاعدة البيانات الجيدة:

تقوم جميع النظم الجديدة بإدخال بعض الأساليب الجديدة لإدخال البيانات ، ومعظم هذه

النظم تحاول اكتشاف تكنولوجيات إضافية للمستقبل. وقد شرح لنا أولارى (O'LEARY, M., 1990) ذلك عندما قال ، إن إدخال البيانات بواسطة لوحة المفاتيح مازال الطريقة السائدة ، ومع ذلك فهناك نظم عديدة قد تحولت إلى الإدخال المباشر بواسطة اختصاصى المعلومات الذى يقوم بالتكشيف والاستخلاص . ولقد كانت المرحلة الأولى لتطبيق نظام الإنتاج الجديد للمستخلصات الكيميائية CABI هو وضع حاسب شخصى على مكتب كل واحد من المستخلصين . وقد وفر ذلك عملية مرهقة خاصة بكتابة المستخلص باليد ثم إدخالها بلوحة المفاتيح ثم تحريرها ومراجعتها وتصحيحها .

أما فى نظام المعلومات الوطنى NTIS فقد طور نظاماً معقداً للغاية حتى يكون قادراً على تناول مختلف الأشكال FORMATS وتزويد مختلف المصالح الحكومية بالمعلومات الإلكترونية ، ويركز النظام الوطنى NTIS على شبكة الإنترنت وغيرها من الوصلات لمصادر المعلومات الحكومية .

أما فى معهد المعلومات العلمية ISI فهم يرون التصوير IMAGING كموجة المستقبل ، وذلك حتى تقدم خدمة إحاطة جارية وتزويد الوثائق والاختزان والتوصيل الإلكترونى للوثائق باعتبارها نشاطاً محورياً للنجاح فى المستقبل ، وإلى جانب تركيز معهد المعلومات العلمية على تكنولوجيا التصوير فهو يهتم بالاتصال بالناشرين الذين يقومون بالنشر الأولى لأخذ الحقوق الإلكترونية منهم ، ووضع نظام قانونى LICENSING SYSTEM بالاتفاق مع الناشرين ويقول O'LEARY أنه حتى عام ١٩٩٠ كان هناك حوالى ٥٧ ٪ من الـ (١٩٩٠) ناشر الذين تم توقيع اتفاقات معهم للحقوق الإلكترونية ، ومن الملاحظ أن معهد المعلومات العلمية لديه (٢٧٨٤) ناشراً للدوريات التى تحتويها منتجاته، كما قام المعهد بعقد توصيل الوثائق مع حوالى ٧٠٪ منهم .

ولكن المستقبل بعد هذا كله مازال غامضاً ، فماذا سيكون عليه نظام توصيل المعلومات المستقبلى ؟ وما دور الناشرين للإنتاج الفكرى الأولى وللإنتاج الفكرى الثانوى (التكشيف والاستخلاص) وخدمات توصيل الوثائق ، خصوصاً والمؤلف سيمكنه النشر مباشرة على الإنترنت ، كما سيقوم العاملون بتوصيل الوثائق بإنشاء قواعد معلومات مجانية لتسويق خدماتهم ، وسينتج الناشرون قائمة محتويات إلكترونية للإنتاج الفكرى الأولى (الدوريات والكتب) ، كما سيبحث الناشرون للإنتاج الفكرى الثانوى فى إعداد النص الكامل للوثائق ، فضلاً عن التكشيف والاستخلاص والوقت الحاضر هو وقت إعادة النظر فى تقييم خدمات التكشيف والاستخلاص الحالية .. ذلك لأنه فى هذا العصر الذى يتميز بتوفر النص الكامل يطرح السؤال التالى ، هل هناك حاجة إلى خدمات التكشيف والاستخلاص ؟ .. ويذهب الباحث أولارى O'Leary إلى أن هذه الخدمات يمكن أن تستمر مادامت ستحدث وتراجع خدماتها بحيث تقدم للمستهلك ما يريد ، أى تقديم نوعية عالية من المنتجات المكشفة على مستوى عال أيضاً ، وتقديم البيانات فى مختلف الأشكال .. فالهيئات التى ستعيش فى المستقبل هى تلك التى ستعرف حاجة السوق جيداً ، وتستطيع الاستجابة السريعة لتجميع المعلومات وتنظيمها وترشيحها للمستفيدين .

٥ - توقعات لانكستر وتنبؤاته :

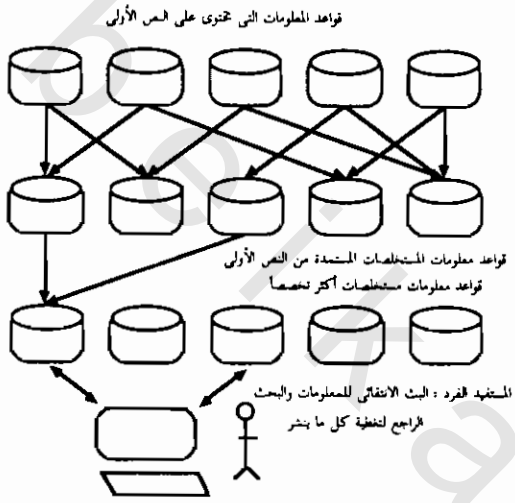
لعلنا نختم هذه التوقعات ببعض التنبؤات التى أوردها لانكستر فى نهاية كتابه عن التكشيف والاستخلاص (Lancaster, F., 1991) وكتابه الأحدث عام (١٩٩٨) بنفس العنوان .

الهجرة من الورق إلى الإلكترونيات وبعض التوقعات المستقبلية :

قام الباحثان لانكستر وجولد هور بإرسال استبيان إلى عدد (٢٠٠) مكتبة فى المصالح الحكومية والصناعة والجامعات ، وذلك لتقدير مدى هذه الهجرة إلى الإلكترونيات .. وقد اكتشفا قيام عديد من هذه الهيئات بإلغاء اشتراكاتهم فى خدمات التكشيف والاستخلاص ، وإن كان ذلك يعزى أيضاً لانخفاض ميزانيات المكتبات وانخفاض القوة الشرائية .. كما اكتشفا أن المكتبات الحديثة الإنشاء تتجه مباشرة إلى الإتاحة على الخط المباشر دون اللجوء إلى المرحلة الانتقالية وهى المصادر المطبوعة .

هذا ويحتوى الإنتاج الفكرى على عديد من التوقعات أو التنبؤات الخاصة بمستقبل خدمات التكشيف والاستخلاص ، ومن بينها توحيد النشر الأولى والثانوى خصوصاً والظروف المعاصرة تشجع على ذلك .. حيث يمكن أن تكون الخدمات الثانوية الببليوجرافية (خدمات الإتاحة) مستمدة آليا من المواد الأولية الإلكترونية ، وأن تكون متاحة فى نفس الوقت الذى يصدر فيه الإنتاج الفكرى الأولى .. وبالتالى فيمكن للمستفيدين الانتقال من أدوات الإتاحة هذه إلى النص الكامل للوثائق على نفس النهاية الطرفية على الخط المباشر.

بل ويتوقع لانكستر الاختفاء الكامل لمعظم – إن لم يكن كل – الخدمات الثانوية ، ومع ذلك فهو يتوقع زيادة الاهتمام بالمستخلصات الإلكترونية (Kuhlen, 1984) كآلية هامة لترشيح الاسترجاع المحسب ونظم البث الانتقائى للمعلومات ولعل الشكل التالى أن يعكس هذا التطور .



مستويات الترشيح فى بيئة غير ورقية
Lancaster, F. (1991), p. 260

خامساً: بعض النتائج والتوصيات :

١ – على الرغم من التناقص التدريجى لبحوث التكشيف والاستخلاص ، إلا أن البحوث النظرية المتصلة بنظم التكشيف والاستخلاص وتطبيقاتها قد زادت تعقيداً خصوصاً بالنسبة للنظم المفهومية Conceptual systems فضلاً عما ورد فى الدراسة عن خريطة البحوث ومجالاتها الخمسة (انظر الجزءين الأول والثانى من الدراسة) .

٢ – التعقد الموضوعى فى مجالات البحوث : لعل المصطلحات الجديدة فى العناوين فى المتن ، والتى لا يتوفر لها ترجمة مقبولة باللغة العربية تعكس هذا التعقد ، فضلاً عن استخدام

الإحصاء والرياضيات والحاسبات فى معالجة هذه الموضوعات المعقدة (انظر الجزء الثالث من الدراسة) .

٣ – فى المنهج والأدوات : يستخدم المنهج التجريبى فى عديد من هذه البحوث ، ومعروف ندرة استخدام هذا المنهج فى دراسات المكتبات والمعلومات بصفة عامة .

٤ – فى بحوث الفريق : معظم هذه البحوث قام بها أكثر من باحث واحد من تخصصات وأقسام علمية مختلفة ، وعلى سبيل المثال لا الحصر (قسم الحاسب الآلى / قسم ديناميكية العلوم والتقنية / قسم نظم المعلومات الإدارية / قسم العلم المعرفى / قسم الرياضيات / ... إلى جانب أقسام المعلومات والمكتبات والاتصال).

٥ – بحوث بين الجامعات داخل الدولة وعبر القارات : الملفت للنظر أن بحوث الفريق هذه تتم فى حالات عديدة من جامعات مختلفة داخل القطر الواحد وبين جامعات فى أقطار مختلفة وعلى سبيل المثال لا الحصر (جامعة أريزونا وجامعة إلينوى بأمريكا) ، (جامعة شيكاغو وجامعة بارلان بإسرائيل وجامعة توركر بفنلندا) ، (معهد الأردن للعلوم والتقنية ، ومعهد إلينوى للتقنية بشيكاغو) ... إلخ .

ولعل أهم التوصيات التى يمكن أن يوردها الباحثان هو ضرورة زيادة وعمق المقررات فى التكشيف والاستخلاص التى تقدمها أقسام المكتبات والمعلومات العربية ، وانتشار وحدة بحوث التكشيف والاستخلاص ضمن مركز بحوث نظم المعلومات بكلية آداب جامعة القاهرة ، يستقطب فيها بعض علماء اللغويات والمعلومات والحاسبات للمعاونة فى استيعاب هذه التطورات الكبيرة فى المجال وإمكانية تطويع اللغة العربية لاستخدامها .

مصادر الدراسة :

- Belkin, N.J. and W.B. Croft (1992) . Information filtering and Information Retrieval : Two sides of the same coin. ACM Commun. 35 (12), 29-38.
- Cardie, C. (1997) . Empirical Methods in Information Extraction. Al Mag. 18 (4), 65-79.
- Cowie, J. and W. Lehert (1996) . Information Extraction ACM Commun 39 (1), 80-91.
- Hahn, V. and Schnattinger, K. (1997). Deep knowledge discovery from Natural language texts, Proceedings of the third International conference on knowledge Discovery and data mining. New Port Beach, CA. p. 175-178.
- Hodges, Julia (2000) . Automated System for the generation of Document Indexes. In : ELIS, V. 66 Suppl. 29, p. 1-19.
- Lancester, F.W. (1998) Indexing and abstracting in theory and practice. 2d ed . Urbana, Champaign.

Structuring and Summarization, Information Processing and Management, Vol. 33 (2) p. 193-207.

- Wiley, D.L. (1994) . Can the traditional abstracting and indexing services survive ? Database, Vol. 17 (6), p. 18-24.

- Milstead, J. (1994) . Needs for Research in Indexing, JASIS, V. 45 (8), p. 577-582.

- O'Leary, M. (1990) . Producing a database : Many choices for data entry. Data Base. V. 38, p. 343-435.

- Salton, G., S ghal, A.; Mitro, M. and Buckly, C. (1997) . Automatic text

obeikandl.com