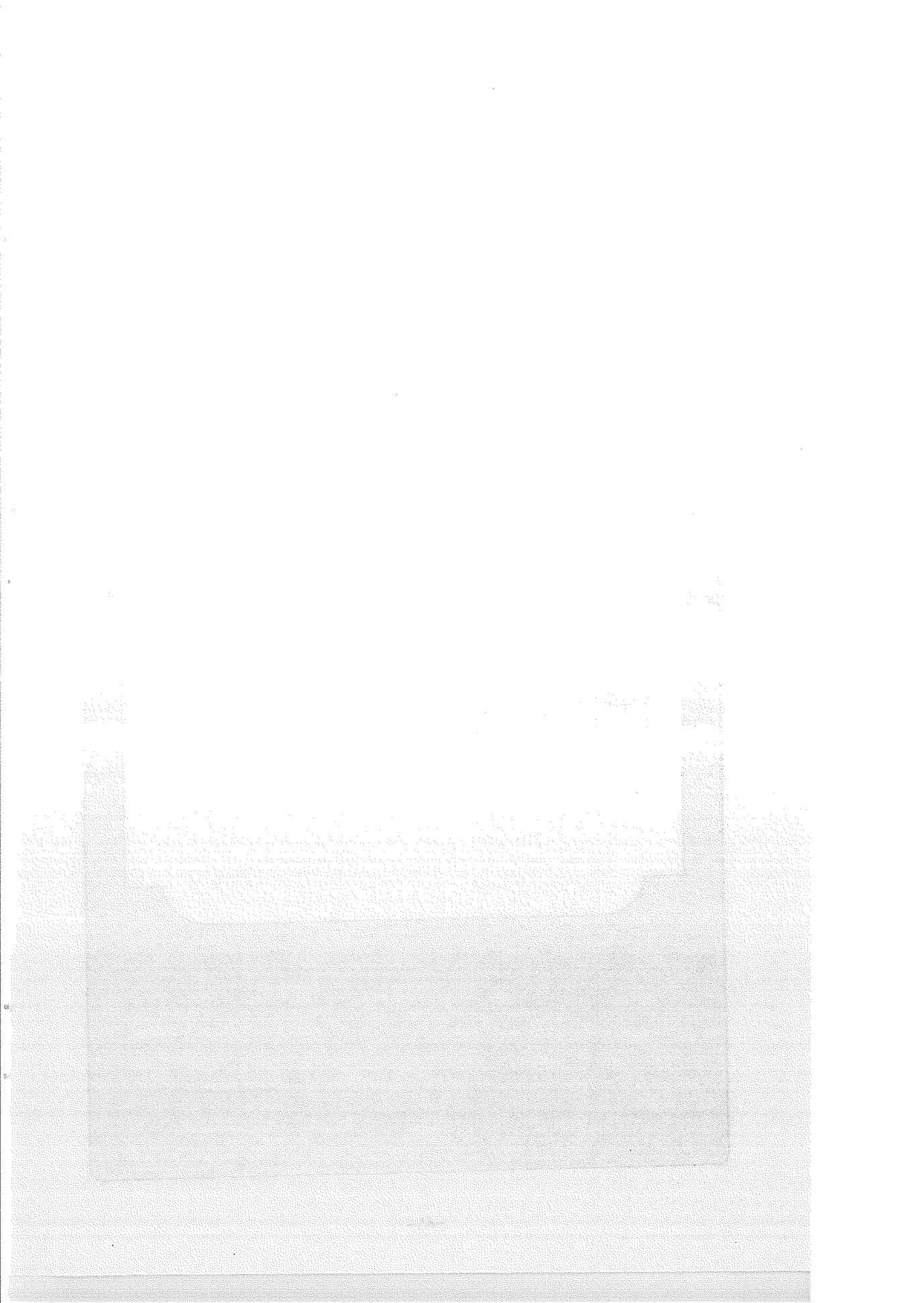
البحث
٢

تصويب الأخطاء اللغوية
للبيانات الببليوجرافية على الحاسوب:

دراسة تطبيقية على قاعدة بيانات ثنائية اللغة.
قائمة المحتويات.

د. أمنية مصطفى صادق

قسم المكتبات
كلية الآداب - جامعة النوفية



مقدمة :

1 – أهمية الفهرس الإلكترونى.

1/1– موقف الإسترجاع الموضوعى فى الفهارس المحسبة.

2 – أبعاد المشكلة و حجمها

1/2 – أبعاد المشكلة.

2/2 – حجم المشكلة.

3 – الإدخال الآلى.

1/3 – أساليب الإدخال الآلى.

2/3 – مراحل الإدخال الآلى.

4 – المراجعة و التصويب بين النظرية و التطبيق.

1/4 – أسلوب المراجعة المطلق.

2/4– إسلوب المراجعة المقارن.

3/4 – تطبيقات التصويب.

4/4 – أدوات المراجعة و التصويب.

5 – أنواع الأخطاء .

1/5 – أخطاء إملائية.

2/5 – أخطاء لغوية.

3/5 – أخطاء الصياغة.

6 – أسلوب خطة التصويب و مراحلها.

مراحل الخطة :

1/6 – مرحلة أولى.

2/6 – مرحلة ثانية.

3/6 – مرحلة ثالثة.

4/6 – مرحلة رابعة.

5/6 – مرحلة خامسة.

6/6 – مرحلة سادسة.

7 – تصويب قاعدة البيانات الببليوجرافية العربية .

8 – تقييم قاعدة البيانات ثنائية اللغة .

9 – الخاتمة.

10 – الإستشهادات المرجعية.

11 – المراجع.

مقدمـــة :

عمدت كثير من المكتبات المتخصصة ذات المجموعات المحدودة نسبيا من حيث الكم و النوع الى إدخال الحاسب الآلى (الشخصى) الى المكتبة ، و تحميل الفهرس البطاقى فى محاولة منها للوصول الى فهرس الكترونى . و ذلك لإستخدامه كبديل عصرى للفهارس البطاقية بأنواعها المختلفة (فهرس الموضوع ، فهرس العنوان ، فهرس المؤلف ، الفهرس المصنف).

و ساعد على إنتشار هذا التطور غير المقنن الإنخفاض الملحوظ فى أسعار الحاسبات الصغيرة فى الآونة الأخيرة ، بالإضافة الى وجود كثير من الا برامج المحلية {1} التى عمد مصمموها على إضافة مداخل إضافية للإسترجاع [كالناشر ، سنة النشر ، الخ ...] إعتقادا منهم أن هذا هو الإستغلال الأمثل لإمكانيات الحاسب و أن مثل هذه الإمكانية سوف تزيد من فاعلية الإسترجاع {2} و حسن إستغلال لإمكانيات الحاسب .و فى حقيقة الأمر أن المقولة التى إشتهرت فى بداية السبعينات عن مدخلات الحاسبات و أنه garbage in garbage out هى فى الحقيقة إذا ما إقترنت بقواعد البيانات الببليوجرافية تصبح garbage in nothing out و هذا يجعلنا أكثر حذرا من ذى قبل.

و عليه فوجود بعض البرامج الدولية [3] بإصداراتها الأولية و البرامج المحلية كلتيهما ، يفتقد الى ادلة إستخدام البرامج ذات المواصفات المتكاملة . فإذا ما توافر دليل الإستخدام للبرامج نجده دليل إستخدام قاصرا ذا إمكانية محدودة في تعريف المستخدم بما مكانية البرنامج أو التطبيق. و هذا القصور يحد من فاعلية إستخدام تلك البرامج ، ناهيك عن عدم توافر بعض الإرشادات الميسرة للمستخدم من اجل إستثمار و تطوير قدرات البرنامج و سواء منها مع إحتياجات المكتبات ، و الوصول الى الإستفادة المثلى من البرنامج أو التطبيق المستخدم .

كل ذلك قد أدى بالضرورة الى وجود فهارس الكترونية على ساحة المكتبات العربية ، ليست على المستوى المرجو منها كما يصعب أن تتم مراقبة إمكانياتها الإسترجاعية أو تقييم أدائها بصفة عامة ،وهذا لكثرة هذه البرامج من جانب و كثرة الأخطاء اللغوية في البيانات البيبليوجرافية التي تم إدخالها من جانب آخر.

و من فحص العديد من الفهارس الإلكترونية بالمكتبات المتواجدة على الساحة العربية [4] تبين لنا وجود أخطاء لغوية للمدخلات تعوق عملية الإسترجاع بشكل واضح . و من هنا بدى لنا أهمية دراسة موضوع "تصحيح الأخطاء اللغوية في قواعد البيانات البيبليوجرافية المحسبة "دراسة مستفيضة مع التطبيق على إحدى القواعد البيبليوجرافية ثائرة اللغة أو أكثر إن أمكن ذلك ، لإجراء المراجعة و التصويب و الخروج بنتائج و مؤشرات من هذه التجربة تثرى الاداء في المكتبات بصفة خاصة و البحث العلمي بصفة عامة.

إن الإهتمام بتصحيح الأخطاء اللغوية للبيانات البيبليوجرافية المخزنة على الحاسوب إنما ينبع في الأصل الى الإهتمام بتصحيح الأخطاء اللغوية للنصوص المدخلة على الحاسب بشكل عام . فتطبيقات معالجة النصوص هي التي دفعت الباحثين الى تصميم البرامج المساعدة في خفض نسبة الأخطاء اللغوية عن طريق إجراء التصويبات بشكل آلي. الا أن هذا التصويب يتم بشكل قاصر لأن البرامج تقوم بعملية مقارنة بين كل كلمة واردة في النص و مضامينها بقائمة المصطلحات اللغوية المعتمدة ، و التي سبق إدخالها الى ذاكرة الحاسوب و التي يطلق عليها تجاوزا "القاموس " فإذا ما وجد لها مقابل تغاضت عن الكلمة المقارنة الى غيرها دون المعنى و إنتقلت الى الكلمة التي تليها لتكرار عملية المقارنة مرة اخرى دون الإهتمام بأشكال الكلمة الواحدة أو تعريفها أو تحديد موقعها من الجملة أو التأكد من الصحة اللغوية بوجه عام و هلم جـــر . أمـــا إذا لـــم تجدكلمـــة متطابقـــة للكلمـــة التـــي هـــي بصـــدد

الفحص أو المراجعة فيقوم الحاسوب بالتنبية الى ذلك فقط ، عن طريق الومضة او بإصدار صوت متقطع للتنبية فقط دون إجراء أى نوع من أنواع التصويبات و فى بعض البرامج يتم عرض قائمة مكونة من عدة مصطلحات قريبة فى تكوين بعض حروف الكلمة التى لا يجد لها مقابل ، و بذلك نستطيع أن نقول أن التصحيح للكلمات بإستخدام الحاسوب إنما هو تصحيح شكلى و ليس جوهريا حيث أنه لا يتطرق للمعنى أو التركيبات اللغوية .من هذا يتبين لنا أهمية التصويبات اللغوية كإجراء ضرورى للفهارس الإلكترونية .

1 - أهمية الفهرس الإلكترونى:

إذا أردنا تعداد مزايا الفهرس الإكترونى للمكتبة فسوف نبدأ من نقطة الإنطلاق الأولى و هى مزايا الفهرس البطاقى ، نعدد مزاياه بالكامل ثم نضيف عليها عشرات المزايا ، مع إستبعاد حجم كبير من المجهود البشرى الذى يتم إنفاقة فى عملية الطباعة على الورق (البطاقات) و مراجعة التنظيم للبطاقات داخل الفهارس و الذى كان يستلزم مجهودا بشريا وتركيزا فوق العادة ، حيث إحتمالات الخطأ الشائع فى الترتيب الألفبائى للبطاقات سواء فى نظام الترتيب بالحرف أو الترتيب بالكلمة من القراء للفهارس.

ثم ننقل الى الإنطلاقة الثانية و هى إمكانيات الحاسوب للفهارس من حيث الدقة فى الترتيب لكل مدخل من المداخل التى يتيحها البرنامج أى الترتيب الألفبائى و السرعة فى النسخ و التكرارية و الإسترجاع هذا بالإضافة الىإمكانيات الإسترجاع الواسعة بالبيانات الأساسية كالمؤلف / العنوان / السلسلة / رقم التصنيف / الموضوع بأمكانيات الجمع و الطرح المنطقى{ 5} هذا بالإضافة الى إمكانيات الإسترجاع بالبيانات الثانوية مثال : سنة النشر / الناشر ...الخ و التى من خلالها يتم عمل إحصاءات شاملة لمقتنيات المكتبة. و هذا يرتبط بامكانيات كل برنامج أوالتطبيق فى عملية الإسترجاع مع خفض تكلفة الصيانة للفهارس و توافر الدقة العالية ، و ضبطها مع إتاحة وسيط الكترونى يمكن من طباعتها بأشكال مختلفة مثال على ذلك : القوائم ، الإشارات البليوجرافية الكاملة ، الإشارات البليوجرافية المختصرة.

تناولت كثير من أدبيات علم المكتبات موضوع أهمية الفهرس الموضوعى و لسنا فى حاجة الى إعادة إثبات أهمية هذا الفهرس و إنه أهم الفهارس على الإطلاق خاصة فى الدول النامية و التى لا يتبوء فيها إسم المؤلف المكانة أو الشهرة التى تسمح بمنافسة الفهرس الموضوعى فى عملية الإسترجاع.

و إذا كان هذا الحال مع الفهارس البطاقية فإن الوضع يزداد أهمية كما سبق توضيحة بالنسبة للفهارس الإلكترونية و التى عمدت الى تطوير نهج الإسترجاع الموضوعى من التركيب القبلى {7} أو السابق Pre-coordination الى التركيب اللاحق {7} post-coordination ليزداد الإسترجاع دقة معتمدا فى ذلك على إمكانيات الحاسب الآلى فى الجمع و الطرح المنطقى و المعروف بإسم الجبر البولينى Boolian algebra و إستلزم ذلك إنشاء المكانز المتخصصة و التى لا يصلح إستخدامها الا من خلال الحاسب الآلى حيث ثبت بالفعل فشل تجربة إستخدام المكانز فى النظم اليدوية {5}.

1\1 – موقف الإسترجاع الموضوعى فى الفهارس المحسبة.

على الرغم من الأهمية القصوى للفهرس الموضوعى اذا ما إرتبط بالحاسب الآلى ، وعلى الرغم من تطور عملية الإسترجاع الموضوعى للمساهمة بطريقة أكثر إيجابية فى تصميم النظم الخبيرة ، الا أن الإهتمام بكفاءة الإسترجاع – فى المكتبات المعنية – لا يتعدى كونة أحد عناصر الإسترجاع دون أى تركيز على التقييم المرحلى فى عملية التحسيب للفهارس بشكل عام . هذا بالإضافة الى عدم إعطاء أهمية علمية لعملية تحسيب الفهارس ، وإعتبار عملية تحميل الفهارس ما هى إلا عملية إدخال آلى أى أن عملية التحول من النظام البطاقى الى النظام الإلكترونى للفهارس ما هى الا عملية إدخال بيانات ببليوجرافية للتخزين على ذاكرة الحاسب الإلكترونى فقط بل و أيضا التغاضى عن كثير من الدقة فى هذه المهمة ظنا من القائمين عليها بأن الحاسب الإلكترونى يمتلك من القدرات التى تسمح بحرية التحليل الموضوعى و أن الفهارس الموضوعية نادرا ماتقوم بتلبية حاجة المستفيد بقدر من الدقة المتوقع لها أو المرجو منها.

من هذا المنطلق ترى الباحثة ضرورة بحث مشكلة عدم كفاءة الفهرس الموضوعى للحاسب الإكترونى و العمل على حل هذه المشكلة بأسلوب عملى و فى حدود الإمكانيات المادية و اللإكترونية المتاحة {8}. حيث أن Bodwin قد أوضح أن تقييم برامج الحاسب إنما يعتمد أساسا على إمكانيات الأجهزة المادية أى Hardware .

2 – أبعاد و حجم المشكلة .

إن خطورة المشكلة ، تكمن فى المخادعة التى تنجم عنها، حيث يتصور أو يعتقد أن مجرد تحميل الفهارس على الحاسب هو تحقيق الإستفادة بكل مميزات الفهرس الإلكترونى و لا يخطر ببالنا ، حجم الأخطاء اللغوية التى تعوق عملية الإسترجاع .فالمكتبة المحسبة تعنى ثقة فى نظام الإسترجاع ،الى درجة كبيرة ، فى نفس الوقت الذى لا يقوم هذا النظام بالمهام الذى إنشىء من أجلها فى أدنى صورة لة و هو الشكل البطاقى.

2\1 – أبعاد المشكلة

للمشكلة بعدان يصعب الفصل بينهما ،بل إن علاقة الربط واضحة. فهناك تداخل تام بين البعدين (الأول) وهو بعد بشرى ،أى ما يتعلق بمهارات المفهرس و مدى المامه ببؤرة التخصص الموضوعى للمكتبة ،بالإضافة الى إجادتة للغة الأجنبية التى يفهرس بها ومدى قدرتة الذهنية على إستيعاب سياسة التحليل الموضوعى بالمكتبةإذا كانت هناك سياسة محددة مسبقا. و الموضوعة بناء على دراسة وافية من أجل تلبية إحتياج المستفيد فى المقام الأول من جانب ، وإمكانيات برنامج الحاسب المستخدم فى تلك المكتبة من جانب آخر. حيث أن إمكانيات الحاسب الآلى ترتبط بالبرامج و التطبيقات المستخدمة بالإضافة الى الإمكانيات المادية للحاسب وقصد بها ال Hardware

أما البعد (الثانى) للمشكلة فهو يتعلق بالأدوات المستخدمة فى التحليل الموضوعى سواء كانت أدوات مطبوعة أو برامج مساعدة و يمكن إجمالها فى الأسباب التالية:

1– المزج بين منهج رؤوس الموضوعات التقليدية Pre-coordinator و منهج الواصفات المستخدم فى المكانز .

2 – تعدد أنواع المكانز المستخدمة و تباينها من حيث التخصص الموضوعى و تباين سنوات النشر الأمر الذى يزيد من مشكلة التشتت اللغوى من حيث التخصص الموضوعى من جانب ومن حيث التطور اللغوى من جانب آخر .

3 – عدم وجود قائمة إسناد مسبقة يعتمد عليها المفهرس و يستد بها .

4 – صعوبة إنشاء شبكات (إحالة) نظرا للتشتت اللفظى و المنطقى الحادث كنتيجة للسبب (1) و (2) السابق ذكرهما.

5 – محدودية إمكانية الإسترجاع الإلكترونية فى مثل هذه البرامج و إقتصارها على الواصفات المفردة أو الكلمة الأولى من الواصفات المركبة و هى صعوبة محدودة يمكن التغلب عليها بإيجاد القواعم التبادلية أو الدوارة Kwic,Kwoc,Kwac أو إستخدام لغات البرمجة المستحدثة و إمكانياتها المتطورة فى البحث عن كلمة فى السياق أو فى جملة.

2\2 – حجم المشكلة.

إن التشتت الواقع لأوعية المعلومات تحت عدد من المداخل الموضوعية دون أن يربطها أى علاقة منطقية ،مع وجود كم من المترادفات بالإضافة الى مشاكل التطور اللغوى و إختلاف مستوى التحليل الموضوعى بين العمومية و التخصص ، مع الإفتقار الى تقنين المستجدات من المصطلحات العلمية قد خلق تضاربا واضحا فى التعبيروهذا قد أدى بالضرورة الى تشتت أقرب ما يكون لضياع منطقية البحث من أجل الإسترجاع.
لذا كان لابد من إجراء دراسة تحليلية لفهرس الكترونى بإحدى المكتبات المتخصصة للوقوف على أبعاد المشكلة ، و لتحجيم المشكلة تم طباعة الكلمات المفتاحية المستخدمة فى الإسترجاع الموضوعى فى تكشيف الأوعية بدون تكرارية أى unique وإحصاء الكلمات المفتاحية الإنجليزية – فى قاعدة البيانات البيليوجرافية{ 6} ثم الوصول الى الأرقام التالية:-

– عدد الأوعية 300, 11 ألف عنوان أجنبى (إنجليزى و فرنسى) .
– مجموع الكلمات المفتاحية المستخدمة فى التكشيف 236, 55 كلمة (مكررة) .
– مجموع الكلمات المفتاحية المستخدمة (بدون تكرار) 656, 7 كلمة.

من هذه اللإحصاءات يتضح لنا حجم المشكلة إذ أن معدل الكلمات المفتاحية المستخدمة فى التكشيف لكل وعاء هو 9و4وهى نسبة مقبولة وهو ما يتناسب مع المؤشرات الموضوعة من قبل {9} عن حجم المكانز المتخصصة و هى 6 آلاف من الواصفات المتخصصة قد أخفت اى شبهة قد توجة الأنظار الى حجم الكلمات المفتاحية المستخدمة فى هذا المضمار.
و علية فقد تم إعادة فحص الكلمات المفتاحية المستخدمة ولكن فى هذه المرة (دون تكرار) unique من خلال وسيط مطبوع فتبين وجود أخطاء متنوعة تصل فى مجموعها الى نسبة 54 ٪ من مجموع الكلمات المفتاحية المستخدمة ، وهى نسبة توضح سبب الشكوى و عدم الدقة فى عملية الإسترجاع من خلال الفهرس الموضوعى حيث يمكن إرجاع الإخفاق فى الإسترجاع الى نفس نسبة الأخطاء الموجودة بالكلمات المفتاحية – أيا كانت نوعية الخطأ الحادث – وهى نسبة غير مقبولة بأى حال من الأحوال .

إن هذا الحجم من الأخطاء قد أدى بنا الى النظر للفهارس نظـرة كلية فى جميع مراحلها ، و عدم الإقتصار على منطقية الإسترجاع Logic of retrieval و علية فقد إستحوذ الإدخال الآلى على قدر كبير من المراقبـة و التحليل للوقوف علـى دور الإدخال فى هذا الحجم من الأخطاء.

أما فيما يتعلق بمجموعة المقتنيات العربية فقد كانت طبيعة الأخطاء تختلف إختلافا جوهريا مما أدى بنا الى إفراد جزء خاص و مستقل عنه و هو الفقرة (7).

3 – الإدخال الآلى:

إن إدارة عملية ميكنة المكتبة أى تحويل الفهارس المطبوعـة بأشكالها المختلفة سواء كانت بطاقية أو غيرها (نموذج إدخال) تعتبر من المهام الشاقة و المشوقة فى آن واحد ، و الخبرة العربية فى هذا المضمار مازالت محدودة تحتـاج الى كثير من التجارب الخاضعة للبحث العلمى حيث أن الإعتماد على التجارب الغربية بكل ما فيها من دقة لا تستطيع أن تضمن سلامة عملية الإدخال بكاملها . إن ملائمة أسلوب العمل و الطريقة المتبعة يجب أن تتخللها نوع من الموائمة لبيئة العمل العربية.
فمازالت الفهرسة بشقيها تتم على الوسيط الورقى (نموذج الإدخال) نظرا لسبب أو أكثر من الأسباب التالية:–

1 – عدم قدرة المفهرس على إستخدام مفاتيح الإدخال الآلى بالسرعة الكافية ، مع عدم الرغبة فى التعلم و إكتساب هذه المهارة.نظرا لإرتباط هذه المهارة بحملة المؤهلات المتوسطة من المتعلمين فى إستخدام الآلة الكاتبة.

2 – عدم توافر البرامج المساعدة فى عملية الإدخال الآلى و التى تختلف فى تصميمها على مستوى شاشة الحاسب عن تلك التى تقوم بعرض البيانات البيليوجرافية حيث تمتاز الأولى بوجود أمكانيات تبادلية بين المفهرس و الحاسب Interactive system

3 – الفصل التام بين العمليات الفنية و بين عملية الإدخال الآلى على مستوى العمل بالمكتبة و يرجع هذا الى إقتصار و ظيفة الإدخال الآلى علـى حاملى المؤهلات المتوسطة و عدم إقبال حاملى المؤهلات العليا على هذا المضمار نظرا لعدم إكتسابهم مهارة إستخدام لوحة مفاتيح الإدخال بسرعة مناسبة للقيام بهذه المهمة بسبب عزوفهم عن هذه المهمة إعتقادا خاطئا ، إعتقادا منهم بأن هذه المهمة لا تليق بحاملى المؤهلات العليا.

4 – رغبة إدارة المكتبة فى الإحتفاظ بالوسيط الورقى كدليل مادى تتمكن من خلالة محاسبة المفهرسين على معدلات الأداء و تحديد مسؤلية العمل فى مراحلة المختلفة.

1/3 – أساليب الإدخال الآلى.

و قبل الحديث عن مراحل الإدخال الالى يجب أن نستعرض أساليب الإدخال الالى المتاحة ، حيث أن نسبة الأخطاء تتوقف على أسلوب عملية الإدخال فى حد ذاتها.

أولا : إسلوب إدخال التسجيلة أو أسلوب الإدخال الذى يمكن تسمية بالإسلوب اليدوى و ذلك نظرا لإن إدخال بيانات البطاقات كل على حدة بطريقة مماثلة تماما لإسلوب كتابة نموذج الإدخال الورقى ، و هو الأسلوب الأكثر إنتشارا و يهدف هذا الأسلوب الى تكوين ملف للبيانات الببليوجرافية يمكن من خلالها تكوين أو إستخراج قوائم إسناد .

و هذا الإسلوب إنما يتطلب جهدا مضاعفا فى عملية الإدخال و المراجعة ناهيك عن تصحيح الأخطاء التى تتطلب وقتا وجهدا ذهنيا بالإضافة الى إمكانيات شخصية فى إجادة التعامل مع الملفات اللإلكترونية و بعض البرامج القصيرة للقيام بهذه المهمة و هو تصحيح الأخطاء التى هى محور هذا البحث.

ثانيا: اسلوب الملفات المفتوحة أو الكشافات الإلكترونية و هو يتطلب دراية و الماما كافيا ببناء قواعد البيانات الببليوجرافية و إمكانية التعامل مع الملفات الإكترونية بمهارة و تمرس واسع حيث يتم إدخال البيانات الى الكشافات مباشرة دون استعانة بشاشة الإدخال المعدة خصيصا للربط بين عدة ملفات كل خاص ببيان ببليوجرافى محدد و لا يربطة بيان بالآخر الا رقم مسلسل.

ثالثا : الإسلوب الثالث و هوالجمع بين بين الأسلوب الأول و الثانى ويكون قاصرا على نوعية معينة من المقتنيات التى يتكرر عدة بيانات فيها بشكل ثابت مثال: الخرائط ، حيث يمكن أن يكون مقياس الرسم ، و حجم ، و نوعية الوسيط : واحدة و ليس بها أى إختلاف . و يمزج بين الإدخال للتسجيلة و الإدخال عن طريق الملفات المفتوحة بحيث يبدأ المفهرس بإدخال البيانات الموحدة أولا ثم يلى ذلك مرحلة إدخال ثانية يتم إستخدام أسلوب الإدخال بالتسجيلة.

رابعا: أسلوب مساعد و لا يمكن الإعتداد به كأسلوب مستقل فى حد ذاتة . وهو إدخال قوائم الإسناد مستقلة مسبقا و هذه القوائم يتم نقلها من الفهارس الطبوعة بشكل رأسى و تخزينها على

الحاسب . ليتم إجراء التصويبات اللازمة بمراحلها المختلفة و إستخدامها فى عملية الإدخال الآلى لتقليل حجم الأخطاءالإملائية و اللغوية الى نسبة الصفر.

2/3 – مراحل الإدخال الآلى.

من مراقبة الواقع الفعلى يمكن تحديد مراحل الإدخال الآلى الى مايلى:
1 – إستلام أوعية المعلومات و الإمضاء بالإستلام على سجل مطبوع.
2 – إتمام عملية الفهرسة و التصنيف على وسيط ورقى.
3 – تسليم أوعية المعلومات و نموذج الفهرسة [مجتمعين] الى مدخل البيانات . وذلك للإستعانة بالوعاء عند الحاجة حيث أن الحروف الأجنبية قد يشوبها عدم الوضوح حين لا يتميز المفهرس بخط واضح.
4 – طبع البيانات الببليوجرافية التى تم إدخالها فى مجمعات تتراوح بين 50 و 100 تسجيلة فى المراجعة الواجدة و تتم الطباعة من خلال الطابعات Dot matrix و التى تتصف بعدم الوضوح.
5 – تتم المراجعة لكل تسجيلة على حدة.
6 – عند وجود خطاء يتم تصحيحة على الوسيط الورقى .
7 – ثم يتم التصحيح النهائى على الوسيط الإلكترونى دون إجراء أى مراجعة نهائية.
من هذا التسلسل يتبين لنا مواطن الضعف الواضحة فى نظام الإدخال الالى ، و المسببة لهذا الحجم الهائل من الأخطاء اللغوية فإذا ما اضيفت اليها أسباب المشكلة و المذكروة فى الفقرة (3) .
نكون قد ألممنا بخيوط المشكلة جميعها ، و يبقى لنا البحث عن أسلوب لتصحيح هذه الأخطاء الفادحة و التى يترتب عليها ضياع أوعية المعلومات و هى على الأرفف لم تغادر أماكنها وتعطى صورة غير صحيحة عن مقتنيات المكتبة.

4 - المراجعة و التصويب بين النظرية و التطبيق.

إهتمت الأبحاث العلمية بنظم المراجعة و التصحيح لقواعد البيانات البيليوجرافية و النصوص اللإلكترونية فمنذ أوائل الستينات اعتبر هذا الإجراء من اهم و أصعب مراحل الإدخال الآلى للبيانات . و مثل هذا النشاط العلمى - أى المراجعة - لا يعتبر و ليد إختراع أجهزة الحاسبات الالية إنما هو نشاط أكاديمى يضرب بجذورة فى تاريخ علوم اللغة . فعملية مراجعة التصحيح اللغوى للأخطاء إنما هى ترتبط بنشأة اللغات ، و من المسلم به أن أدوات المراجعة و التصيحح مثل القواميس اللغوية بكافة أنواعها و أحجامها ، هى الشاهد و الدليل على ذلك و هذا بالإضافة الى أنها المرجع الأخير اذا إختلفت الاراء على صحة كلمة أو تعبير .

لقد عمدت شركات برامج الحاسب الإلكترونى الى تحميل جزئى للقواميس اللغوية على الوسيط الإلكترونى و يقصد بالتحميل الجزئى هو إدخال أجزاء من القاموس المطبوع الى الحاسب . هذه الأجزاء نعنى بها أنه تم إستبعاد كثير من الإشتقاقات اللغوية [10]، و ذلك حتى تقوم تلك القواميس بدور المصحح فى مرحلة المراجعة للنصوص اللغوية المدخلة حديثا لملفات قواعد البيانات المختلفة ، دون البيليوجرافية، فى محاولة لتقليل نسبة الأخطاء اللغوية.

وعلى الرغم من الآمال التى كانت معقودة على مثل هذه القواميس إلا أنها أى - القواميس الإلكترونية - لم تحقق نجاحا ملحوظا إلا فى حدود ضيقة و هى حدود معالجة النصوص العامة كالخطابات و التقارير .
و يرجع هذا القصور أيضا الى طبيعة اللغات الحية و ما تحكمها من قواعد لغوية و تصريفات لأفعالها ، و مثل هذه القواميس لا تلعب دورا ملحوظا فى قواعد البيانات البيليوجرافية لما لهذه القواعد من طبيعة خاصة تختلف إختلافا بينا عن النصوص العامة.
فقواعد البيانات البيليوجرافية تخضع لصفة التخصص الموضوعى و ما تتسم به هذه التخصصات من تطور ديناميكى فى المصطلحات المستخدمة مما يقلل من فاعلية القواميس اللغوية العامة ، أى الغير متخصصة فى فرع محدد من فروع المعرفة نظرا لما يطرأ على لغة التخصص من إصطلاحات ذات معنى و مدلول محدد مع توقف إستخدام بعض المصطلحات الأخرى مثال على ذلك مصطلح : الإتحاد السوفيتى.

ولكن و على الرغم من ذلك يمكننا أن نعتبر أن قوائم الإسناد بأنواعها المختلفة قائمة الإسناد الموضوعى ، قائمة الإسناد بأسماء الناشرين العالمية هى بمثابة القاموس الخاص الذى يمكن الإعتماد علية فى قواعد البيانات البيليوجرافية سواء كانت محمولة على و سيط ورقى او الكترونى . و لكن، قائمة الإسناد ،هذه فى حاجة إلى من ينشئها فى البداية و تخضع عملية الإنشاء، الى عملية ادخال آلى وذلك يحتاج بالضرورة الى مراجعة و تصويب . من هنا يتبين لنا أهمية الإلمام بأسلوب المراجعة و التصحيح بأبعادة المختلفة ليس فقط لإستخدامة فى إنشاء قواعد البيانات البيليوجرافية فقط و لكن أيضا فى إعداد الأدوات المساعدة مثل قوائم الإسناد و قواميس المراجعة و التى يجب أن تتصف بدقة عالية ، حتى تكون ذات فاعلية خاصة اذا ما كانت محملة على وسيط الكترونى، و ذلك لإن إمكانيات البرامج تعتمد فى هذه الحالة على الكلمة أو الإسم أو السياق المراد نقلة بطريقة الإشارة الضوئية Highlight أى النقل التصويرى أى يكون النقل صورة طبق الأصل لما تم الإشارة الية سواء كان به أخطاء أو لم يكن به ، أى بدون أى تدخل بأى شكل من الأشكال فى عملية الإدخال.

إن التجارب العديدة فى مجال المراجعة و التصويب قد افرزت العديد من النظريات العملية التى تبناها العلماء فى محاولات إثبات أهميتها أو العكس أما فيما يتعلق بنظام مراجعة قاعدة البيانات البيليوجرافية فيخضع الى أسلوبين كما شرحهما Pollock & Zamora فى مقالهما عام 1983 و هما:

1/4 - إسلوب المراجعة المطلق.

و هذا الإسلوب يمكن أن نطلق علية أسلوب المراجعة 'الحر' حيث يعتمد المراجع فى المرتبة الأولى على الذاكرة الحية فى الحكم على صحة كل كلمة مفردة أى من واقع الذاكرة الشخصية أو بالرجوع الى القواميس و المراجع إذا إستلزم الأمر و يقصد بصحة الكلمة كل ما يتعلق بالتركيب اللغوى و الصياغة و الموقع من الكلمات الأخرى.

2/4 - إسلوب المراجعة المقارن.

أما أسلوب المراجعة المقارن فهو أكثر تعقيدا عند شرح خطوات تنفيذه خاصة و هو بأختصار إعتماد القائم على المراجعة على مقارنة الكلمة بما يسبقها أو ما يتبعها من كلمات و هذا الإسلوب هو ما لجأت الية نظريات التصويب الإلكترونية فى إعتمادها على القواميس الجزئية ، وهو مقارنة الكلمة المراد مراجعتها بنظائرها فى القاموس الإلكترونى .و يعيب هذا الإسلوب أنه يعتمد فى تنفيذه على عدة مراحل و لكنه من ناحيةأخرى يتسم بالسهولة و اليسر من حيث التنفيذ.

وهذه السهولة تعمل على توفير الوقت بشكل واضح مع المحافظة على القوة الذهنية للمراجع بحيث يمكن له مراجعة أضعاف حجم الكلمات المفتاحية بالإسلوب المطلق أو الحر دون الوصول الى درجة من الإرهاق الذهنى فى وقت قصير نسبيا.

3/4 – تطبيقات التصويب.

لكل أسلوب مما سبق ذكرة طريقة فى التطبيق فنجد أن (الإسلوب الحر) هو أسلوب يعتمد فى المرتبة الأولى على الذهن البشرى و ما لدية من الماما تام بالغة المستخدمة و مدى الماما المصحح بالتعبيرات اللغوية المختلفة و بالتالى فهو فى حاجة الى خبرة طويلةيجب ا ن يتصف بها المراجع فإذا ما كانت المراجعة فى علم متخصص فيجب الإلمام بالتخصص العلمى أولا ثم خبرة المراجعة و التصحيح ثانيا و جدير بالذكر أنه عند الإلمام بالعلم أو بالتخصص فإننا كثيرا ما نقف أمام أفراد على دراية بالتخصص دون الإهتمام بالمصطلحات المستخدمة و مدلولاتها فى هذا التخصص .
أما الإسلوب المقارن فله صعوبات و هى كثرة عدد المراحل التى يمر بها بشكل يصبح معةإدارة عملية المراجعة و التصويب مهمة شاقة فى حد ذاتها.

4/4 – أدوات المراجعة و التصويب.

للمراجعة ادوات لا غنى عنها و هى القواميس و الأدلة المتخصصة، و لكن عند مراجعة القواعد البيلوجرافية ففى كثير من الحيان قد يستلزم الأمر الرجوع الى الوعاء المفهرس للتأكد من لفظ أو من إسم أو عنوان.

1/4/4 – إن أدوات المراجعة لا تقف عند القواميس و الأدلة بل تتعداها الى الوسيط الذى تتم المراجعة علية ، حيث ثبت بالفعل أن عملية المراجعة لقواعد البيانات المحسبة تتم من خلال الوسيط الورقى فى كثير من المراحل لعدة أسباب نذكر منها :-
1 – إعتياد الفرد التدقيق على الوسيط الورقى .
2 – الشعور السريع بالإرهاق لحاسة النظر، إذا ما تم التدقيق من أجل المراجعة على شاشات الحاسبات .
3 – الرغبة فى الإحتفاظ بسجل الأخطاء و سجل التصويبات يمكن الرجوع اليه عند الحاجة الى ذلك .

2/4/4 – الأقلام الملونة و الرموز المستخدمة من الأدوات التى يستعان بها فى عملية المراجعة ، فإستخدام الأقلام ذات الألوان المختلفة يهدف إستخدامها الى تحديد نوعية الخطاء مما يسهل من تصويبة فى المرحلة المخصصة له ، فهو إذا يحدد و يصنف الخطأ فى آن واحد، كما يستخدم إختلاف الألوان فى تحديد إختلاف وجهات النظر أو الرأى أو توحيد الآراء للوصول الى قاعدة موحدة.

5 – أنواع الأخطاء .

تتطلب عملية التصويب وضع خطة محكمة تتصف بالواقعية حتى تكون قابلة للتنفيذ و تصبح فعالة فى رفع مستوى الإسترجاع ، و لكن قبل وضع الخطة هناك ضرورة ملحة للتعرف على الأخطاء الموجودة بقاعدة البيانات البيليوجرافية مجال البحث الدراسة.

هذا و قد تم التعرف على الآخطاء من خلال الخطوات التالية:
1 – التعرف على حجم و نوعية الأخطاء و حجمها و ذلك قبل تحديد وسيلة و أسلوب التصحيح من خلال طبع قائمة بالمصطلحات (مفردة) بالإضافة الى إحصائية بعدد إستخدامات كل مصطلح فى ذلك الوقت ، أى وقت إجراء التصويبات.

2 – توحيد الأخطاء و هو إجراء أولى فى عملية التصويب و خطوة إيجابية على درب التصحيح الشامل حيث تتضح الرؤية من عملية توحيد أخطاء الصياغة و يصبح من السهل وضع خطة يسهل تنفيذها .

3 – تصحيح الأخطاء الإملائية و اللغوية و هو ما سوف يتم شرحة فى الفقرة 6/ 1-6/6.

4 – قد يستلزم الأمر– وهو كذلك فى أغلب الأحيان – الى إعادة صياغة المصطلح المستخدم خاصة إذا ما كان من رؤوس الموضوعات بالصيغة التقليدية Postcoordinator

1/5 – أخطاء إملائية
مثل هذه الأخطاء يمكن تقسيمها الى أنواع مختلفة ترجع فى ذلك الى سبب حدوثها الفئة الأولى و هى الأخطاء الناتجة عن الإسراع المبالغ فية عند الإدخال.

1 – إدخال حرف قبل حرف آخر : Copmuter
2 – إدخال حرف بدلا من حرف : conputer
3 – إسقاط أحد الحروف من الكلمة : coputer

2/5 – و هى الأخطاء الناتجة عن عدم الإلمام باللغة الإنجليزية :
1 – أخطاء الجهل بالكلمة و الإعتماد على الذاكرة دون الرجوع الى النص فى محاولة الإسراع بعملية الإدخال computar
2 – أخطاء الجهل بالنطق السليم و أثر ذلك على الإدخال :combuter

3/5 – أخطاء الصياغة :
1 – الأخطاء التركيبية للواصفات :
وأغلبها يتعلق بمشكلة الصياغة التقليدي و صياغة الواصفات
مثال : <Software and Hardware><Software> <Hardware>
مثال : مصطلح [مجتمعات مدنية] بدلا من <مجتمعات حضرية>

2 – أخطاء منهج التحليل:

– خطأ مستوى التحليل الموضوعى:
– خطأ المفهوم :
– شرك التطور اللغوى دون الربط أو الإحالة: إستخدام مصطلح

3 –المرادفات الشكلية:
و هى إستخدام المترادفات و عدم الإلتزام بالقواعد العامة لرؤوس الموضوعات أو الواصفات . و بذلك تتنافى اللغة المقننة فى عملية الإسترجاع مما يجعل التشتت واقعا ملموسا.

6 – أسلوب خطة التصويب و مراحلها.

من نوعية الأخطاء و جدنا ضرورة تقسيم خطة التصحيح الى مراحل منفصلة كل منها يعتبر مرحلة مستقلة بذاتها و ذلك من أجل :

1 – تسهيل عملية التصحيح بالتركيز على نوعية واحدة أو أثنتين على أكثر تقدير، فى المرحلة الواحدة .

2 – الإسراع فى رفع مستوى الأداء لقاعدة البيانات البيليوجرافية بنسب متفاوته فى كل مرحلة مستقلة عن الأخرى ، و عدم الإنتظار الى نهاية مشروع التصحيح لرفع مستوى الإسترجاع فى قاعدة البيانات البيليوجرافية بالنسبة لما هو مقرر فى بداية الخطة و التى تستغرق عدة شهور لأكثر من سبب مثل إنشغال العاملين بالإدخال الآلى لما يستجد من أوعية جديدة الخ..
هذا و جدير بالذكر أنه من الممكن وضع خطة أخرى تعتمد على تصحيح جزئى أى عدد محدد من التسجيلات فى كل مرحلة . و لكن مثل هذه الطريقة تحتاج الى فصل الكلمات المفتاحية لعدد التسجيلات المحدد فى كل مرحلة على حدة حيث يتم خلالها التصحيح مثال : (تسجيلات من رقم 4000 الى رقم 5000) وتعتمد عملية التصويب هنا على الذاكرة فى تصويب الخطاء و هو مرفوض حيث يمكن أن ينتج عنه أخطاء أكثر تعقيدا).

مراحل الخطة :

1/6 – المرحلة الأولى:
إذا أردنا توخى الدقة فإن العمل فى المرحلة الأولى يرتبط بالفترة التى تسبق وضع خطة العمل و ذلك نظرا للإجراءات التالية:
– طباعة الكلمات المفتاحية فى قائمة مرتبة الفبائية بدون تكرارية للمواصفات.
– فحص عينة مما تم طباعتة بنسبة 10 فى ٪ من الكلمات المفتاحية .
– نتج عن إجراء الفحص تحديد نسبة أخطاء متنوعة مقدارها 53 فى ٪ من عدد الكلمات المفتاحية.

2/6 – المرحلة الثانية:
تشمل هذه المرحلة على :
– تصويب الأخطاء الإملائية : على الوسيط الورقى أولا .
– تصويب الإخطاء الإملائية على نسخة من الملفات الإلكترونية بعد فصلها عن القاعدة الأساسية و ذلك كإجراء و قائى لعدم تلف الملفات الرئيسية ن أو ضياع كل أو جزء من البيانات .
– طباعة القائمة مرة ثانية إستعدادا للمرحلة التالية.
– و لكن نظرا لإنخفاض الكلمات المفتاحية الإنجليزية بعد التصويب من 7427 كلمة الى 5696 كلمة أى تلافى ما يقرب من 1728 كلمة و هذا يعنى نسبة أخطاء تصل الى 25 ٪ من مجموع الكلمات فقد تم إجراء ملحق للمرحلة الأولى يتم فيه تلافى الأخطاء التى سقطت سهوا فى هذه المرحلة ثم طباعة الكلمات مرة أخرى إستعدادا للمرحلة التالية.

3/6 – المرحلة الثالثة:
و تشمل تصويب الأخطاء اللغوية ، على الوسيط الورقى أولا . ثم تصويب الأخطاء اللغوية على نسخة من الملفات الإلكترونية بعد فصلها عن القاعدة الأساسية.
– عمل ملحق للمرحلة الثالثة لتلافى الأخطاء التى لم يتم تصويبها و سقطت سهوا .
– طباعة قائمة بالكلمات الدالة إستعدادا للمرحلة التالية.

4/6 – المرحلة الرابعة:
تحديد الكلمات المستخدمة من القوائم المعتمدة كل على حدة و لقد كانت فى تجربتنا هذه القوائم و المكانز التالية:

- L.C. Subject headings list.
- AID Thesaurus.
- SPINES Thesaurus.

– تم تحديد الواصفات التى تحتاجها المكتبة بصفة خاصة و غير متوفرة فى هذه المكانز أو القوائم سابقة الذكر من حيث المعنى أو المدلول و قد وصلت نسبتها الى 9٪.

– إن حجم و نوعية الكلمات المستخدمة دون إعتماد المكانز السابقة ، قد وصل الى 31 ٪ من مجموع الكلمات المستخدمة مما آثارقضية هى : مدى فهم إخصائى المعلومات لنظام الإسترجاع و إمكانيات البرنامج المستخدم على الحاسوب .

– و بذلك يكون هناك 60 ٪ من الواصفات المستخدة بها مرادفات على اثر إستخدام العديد من المكانز فى آن واحد دون محاولة تجنب المرادفات.

– تحديد إحالة (إنظر) ووضع قائمة مستقلة بهذه الإحالات ، تستخدم كأداة من أدوات التكشيف و يمكن تحميلها على وسيط الكترونى لتصبح قائمة إسناد فيما بعد .

5/6 –المرحلة الخامسة:
و هى مرحلة تنفيذية للمرحلة السابقة أى المرحلة الرابعة على الوسيط الإلكترونى و تشمل إحدى عشرة آلاف تسجيلة .

6/6 - ا لمرحلة السادسة:
و هى مرحلة يتم فيها إعادة المراحل السابقة بتفاصيلها على العدد المحدد من التسجيلات الذى تم إدخاله بالفعل لقاعدة البيانات البيليوجرافية فى الفترة من بدء إجراء عملية التصحيح و حتى نهايتها و التى إستغرقت أكثر من 18 شهرافى الفترة من يناير 1992وحتى يونية 1993.

عند هذا الحد يجب الفصل التام فى الحديث بين الخطوات العملية التى إتخذت فى وضع خطة التصويب الخاصة بقاعدة البيانات البيليوجرافية الإنجليزية و نظيرتها العربية .إن ما يتبادر للذهن للوهلة الأولى هو أن مراحل التصويب المستخدمة لكليهما واحدة و لكن عندما ننتقل الى مرحلة التنفيذ نجد أن هناك خلافا فى طبيعة الأخطاء يستلزم تغيير مراحل الخطة بشكل جذرى و جلية فسوف نبدأ بعرض خطة و مراحل تصويب قاعدة البيانات البيليوجرافية الإنجليزية ثم نستطرد فنوضح الفروق الناجمة فى الخطة و إجراءات التصويب لقاعدة البيانات العربية.

7 - تصويب قاعدة البيانات البيليوجرافية العربية.

نظرا للإختلاف الواضح بين خطة و أسلوب العمل فى كل من القاعدتين العربية و الإنجليزية فقد آثرنا سرد الإختلافات التى طراءت علىخطة التصويب للقاعدة العربية، مستقلة و غير مدمجة .إحتوت قاعدة البيانات البيليوجرافية على أكثر من 4 آلاف تسجيلة تم تكشيفها بما يقرب من 24 ألف كلمة مفتاحية ، و حيث أن سياسة التكشيف لم تكن محددة الإسلوب فى التحليل الموضوعى من حيث السياسة العامة المستخدمة فى تحديد المستوى الموضوعى Level of analysis بالإضافة الى الخلط بين الأدوات كإستخدام رؤوس الموضوعات التقليدية و الواصفات المستقاة من المكانز المتخصصة المختلفة.الأ مر الذى أسفر عنه إنخفاض واضح فى معدلات الإسترجاع لما تحتوية المكتبة بالفعل . فقد تم دراسة الموقف أولا ثم تم وضع خطة تختلف فى بعض مراحلها عن الخطة التى سبق تنفيذها فى قاعدة البيانات البيليوجرافية الإنجليزية . لذا لزم إستعراض الفروق فى تلك المراحل :

أولا : تم حذف حرف التعريف أل و الذى أستخدم بشكل مبالغ فية وصل الى وجود 68 ٪ من مجموع الكلمات المفتاحية تسبقها حرف التعريف أل و التى تسببت فى وجود عبء على النظام و التأثير على سرعة الإداء للأجهزة بالإضافة الى ا لتشتيت فى عملية الإسترجاع .

هذا و جدير بالذكر أنة يمكن إستخدام إمكانيات لغات البرمجة {13} و قواعد البيانات الحديثة و التى توفر (مصفاه) filter يمكن إستخدامها فى إهمال حرفى (ال) التعريف و من ثم ترتيب الواصفات ترتيبا الفبائيا فى نسق مسلسل واحد جيد و لكن هذا لا يمنع ضرورة حذف

الإستخدام المبالغ فيه لحروف التعريف كما إن هناك خطورة فى إستخدام طريقة المصفاة بطريقة عشوائية مع اللغة العربية نظرا لوجود كثير من الألفاظ التى يفضل إحتساب (ال)بها صحيحة (الومنيوم).

ثانيا: تم ضبط الهمزات بكافة أنواعها و ذلك محافظة على سياق اللغة و المعنى .مثال:آ ، أ ، لآ ، إ ،ئ . كما تم ضبط التاء المربوطة و الهاء .

ثالثا : تم توحيد أسماء الهيئات و المجالس و البنوك و الأجهزة الحكومية. مثال : الهيئة العربية للتصنيع ، البنك المركزى .

رابعا : تم توحيد بعض المسميات الأجنبية مثل : الومنيوم ، يورانيوم ، سليكون.

خامسا : تم توحيد مصطلحات الحاسب الآلى و مسميات لغاتة المختلفة مثال ذلك : البيسيك ، فورتران ، باسكال، و جدير بالذكر أن اللغة الواحدة كانت تكتب بأكثر من ثلاثة أشكال مما يعوق عملية الإسترجاع بشكل واضح ، مثال على ذلك : البيسيك ، لغة البيسيك ، باسيك ، بيسيك .

سادسا: تم تحويل رؤوس الموضوعات التقليدية الى واصفات .و ذلك بإتباع الخطوات التالية:-

- عدم قلب رؤوس الموضوعات مثال : علم الإجتماع و ليس الإجتماع ، علم .
- معاملة الأسماء الجغرافية معاملة الواصفات.
- معاملة جميع التفريعات لرؤوس الموضوعات التقليدية على إنها واصفات .

بما فى ذلك التفريعات الشكلية و الوجهية. مثال : طرق البحث ، أسعار ، تدريب.

سابعا : تفضيل صيغة الجمع فى المصطلحات مثل : مجتمع و مجتمعات- محافظة و محافظات . و علية يتم تكشيف الوعاء بإسم المحافظة مباشرة مثال .<بنى سويف>و يضاف مصطلح .<.محافظات>,واصف إضافى لجمع المحافظات كلها عند إجراء إسترجاع يتطلب محافظات الجمهورية جمعاء.

ثامنا : إقرار بعض المصطلحات مثل: <تمويل التعليم >, بدلا من [تكلفة التعليم]وإستخدام مصطلح <إقتصاديات التعليم >{ 14} و إستخدام مصطلح <موارد إقتصادية >بدلا من [جغرافيا إقتصادية] .

كم تم إيضافة <موارد مائية >بدلا من [مصادر مائية] .

و تم إضافة مصطلح < مجتمعات عمرانية جديدة >.

8 – تقييم قاعدة البيانات الببليوجرافيةالتى تم تصويبها:

إن تقييم قاعدة البيانات الببليوجرافية إنما ينقسم الى شقين الأول وهو تقييم صحة البيانات و خلوها من الأخطاء و الثانى و هو تقييم البرنامج المستخدم فى خزن و إسترجاع البيانات الببليوجرافية ، و الذى حفلت به الدوريات التجارية و خاصة تلك التى تتبنى الدعاية للبرامج من خلال إجراء المقارنات بينها من حيث التشابة و الإختلاف ، و علينا الا نغفل أن تقييم أى برنامج إنما يحتاج الى كم من البيانات الصحيحة لإجراء إختبارات السرعة و الأداء و هذا ما سعى الية هذا البحث و هو توفير نظام دقيق و خطوات محددة لتصويب الأخطاء فى قواعد البيانات الببليوجرافية.

بقى علينا إستعراض البيانات الببليوجرافية فى هذا البحث قبل إستخدامها فى عملية الإسترجاع و ليس لنا فى هذا المضمار من دليل قاطع سوى الأرقام حتى تقوم بعملية التقييم ، بفحص ملف الكلمات المفتاحية وجد أن نسبة الأخطاء لا تتعدى 6 ٪ و هو ما يترك قاعدة البيانات بأكثر من 94 ٪ من الكلمات خالية من الأخطاء .الأمر الذى يجعل من مكتبة مركز المعلومات و دعم إتخاذ القرار نموذجا يحتذى به فى مجال المكتبات المتخصصة على الساحة العربية التى تفتقر الى حد كبير الى هذه النوعية من المكتبات رغم أهميتها.

9 – الخاتمة .

و نستخلص مما سبق عرضة و تحليلة فى هذا البحث أنه ، قبل البدء فى تقييم {2} برامج أو تطبيقات المكتبات على الساحة العربية يجب التأكد من صحة البيانات البيليوجرافية المدخلة من خلال خطة محكمة للمراجعة و بالتالى وضع خطة مناسبة لتصحيح الأخطاء.حيث أن تصويب ملف الكلمات المفتاحية يرفع من معدلات الإسترجاع الموضوعى ، ويساعد بالتالى على ا لعمل فى تصويب ملفات أخرى كملف العنوان و ملف المؤلف..الخ و بالتالى خفض نسبة الأخطاء فى الفهارس الإلكترونية الى أقل نسبة ممكنة بحيث تصبح قادرة على إعطاء صورة واضحة عن مقتنيات المكتبة . هذا ومن المعروف أن بعض البرامج التى تعد فى الوقت الحاضر تحاول الإستفادة من أمكانيات تطبيقات النوافذ فى توفير قائمة إسناد الكترونية تساعد كل من المفهرس على ضبط الإسترجاع اللغوى ، و المستفيد على رفع معدلات الإسترجاع بشكل واضح ومن ثم رفع مستوى الآداء فى المكتبات سواء كانت مكتبات متخصصة أو مكتبات عامة و التى يجب أن يتوافر فى كل منهما المحيط المناسب للعمل من تجهيزات و أثاث و إضاءة و غير ذلك من الأشياء التى توفر الراحة النفسية و البدنية للعاملين فى المكتبات و فى نظام خدمات المعلومات لما لها من أهمية كبيرة و تأثير مباشر عن حسن و سلامة الآداء وتوفيرا للوقت ذلك العنصر الفريد و الحيوى لعمليات التنمية بصفة عامة و لخدمات المعلومات بصفة خاصة و الذى لايمكن تعويضة أو إحلالة أو شراؤة أو تخزينة أو زيادتة كغيرة من العناصر و الأجهزة و الأدوات المستخدمة فى المكتبات و مراكز خدمات المعلومات .

الإستشهادات المرجعية:

****

{1} تعتمد غالبية هذه البرامج على تطبيقات Dbase III,or IV

{2} التقييم التجارى يتم من خلال العناصر التالية:-

1- أجزاء وأوجة البرنامج -Features

2- إمكانيات التخزين -Storages

3- إمكانيات الآداء -Performance

4- سهولة الإستخدام -Ease of use

5- أدلة التشغيل -Documentation

6- الدعم الفنى(التدريب، الصيانة...) -Support

7- الآداء العام للبرنامج -Overall performance

{3} لقد تم إستبعاد المكتبات التى طبق فيها نظام CDS-ISIS version 1.0,2.0 و ذلك لإقتصارإمكانية إستخدام البرنامج الأخيرعلى الحاسبات الكبيرفقط Dobbis-Libbis

{4} فهارس الكترونية لمكتبات على الساحة العربية.

{5} Boolean Algebra

{6} مكتبة مركز المعلومات و دعم إتخاذ القرار بمجلس الوزراء – القاهرة.

{ 7} لتركيب القبلى او السابق لرؤوس الموضوعات و التركيب اللآحق للواصفات

{8} يقصد بالإمكانيات المادية و الإلكترونية هو محدودية الطاقة الفنية المتوفرة للتحليل الموضوعى بالإضافة الى محدودية إمكانيات برامج الإسترجاع محليا بمقارنتها بما هو متاح فى الأسواق العالمية و نذكر منها على سبيل المثال لا الحصر In-magic plus GEAC,SoutronTalis, علما بأن البرامج ذات الشهرة العالمية مازالت تواجة مشكلة التعريب بالإضافة الى إرتفاع أسعار الإقتناء و الصيانة نظرا لعدم وجود وكلاء معتمدين لها فى الأسواق المحلية هذا بالإضافة الى إعتماد تشغيلها على أجهزة ذات كفاءة خاصة من حيث سعة التخزين الدائمة و المؤقتة و سرعة المعالجة.

{9} إنظر Lancaster , F.W. (1972) p.73

{10} ظهر حديثا بعض برامج معالجة الكلمات و بها قواميس تحتوى على إشتقاقات الكلمة المشروحة و مثل هذه البرامج قاصرة على معالجة الكلمات : مثل برنامج أبجدهوز لنظام الأبل ماكنتوش.

{11} لقد غستخدم فى هذا البحث الإصطلاات الثلاثة: كلمات مفتاحية ، و هى الكلمات المستخرجة من نص عنواء الوعاء، إصطلاح : كلمات ، و هى الكلمات المؤلفة من مدخل البيانات دون سند من القوائم المعتمدة أو المكانز ، و أخيرا الواصفات و هى Descriptors منتقاه من المكانز أو مضافة و قد روعيت قواعد الصياغة.

{12} مكتبة مركز المعلومات و دعم إتخاذ القرار – القاهرة.

{13} كما يمكن أيضا إستخدام إمكانيا ت الإسترجاع (القاطع) Trancation أو إمكانيات الحرف النكرة Wild caracter

{14} كما جاء مشروحا فى : إقتصاديات التعليم / محمد محروس إسماعيل .– القاهرة : الأهرام .– 1993 (الكتاب الإقتصادى 1993) ص 6.

المراجع :
******

شعبان عبد العزيز خليفة
قائمة رؤوس الموضوعات العربية الكبرى / شعبان عبد العزيز خليفة و محمد حوض العايدى .- الرياض : دار المريخ . 1985 .2مج

محمد محروس إسماعيل
إقتصاديات التعليم .- القاهرة : مؤسسة الأهرام 1993(الكتاب الإقتصادى للأهرام العد د 67-) .

References:
****

Angell, Richard C. & (et.al.)
Automatic spelling correction using a trigram similarity meaure / by Richard C. Angell . George E. Freund & Peter willet.- Information Processing & Management Vol.19,No.4,Pp.255-261,1983.

Berghel ,H.L.
A logical framework for the correction of spelling errors in electronic documents . Information Processing and Management , 23, 1987, 477- 494.

Damerau, Fred J.
A technique for computer detection and correction of spelling errors .- Communication of the ACM Vol.7, No.3, March 64.

Lancaster, F.W.
Vocabulary control for information retrieval . -Washington , D.C.: Information Resources Press .1972.

Mitton , Roger
Spelling checkers , spelling correctors and the misspellings of poor spellers.- Information Procesing and Management Vol.23, No.5, Pp. 495-505, 1987.

Peterson ,J.L.
Computer programs for detecting and correcting spelling errors .- Communication of the ACM Vol.23,No.12 Dec. 80 Pp.676-687.

Peterson ,J.L.
A note on undetecting typing errors .- Communications of ACM, Vol.29,1986,Pp. 633-637.

Pollock ,J.J.
Spelling error detection and correction y computer : some notes and a bibliography .- Journal of Documentation,38,1982,Pp. 282-291.

Pollock, J.J. and Zamora , A.
Collection and characterization of spelling errors in scientific and scholarty text .- JASIS,Vol.34 (1) ;1983,Pp. 51-58.

Pollock, J.J. and Zamora , A.
Automatic spelling correction in scienific and scholarly text .- Communication of the ACM .- Vol.27 ,No.4, April 84.

Sadek , O.M.
Social Science Thesauri ; A quantative & qualitative assessement of six selected thesauri .- (M.A. Thesis) University of Sheffield . 1983.- 129p.

Rogers, Heather J. & Willett , Peter
Searching for historical word forms in text Databases using spelling correction methods: Reverse error and phonetic coding methods .- Journal of documentation Vol. 47, No. 4 .-Dec. 91, Pp. 333- 353.

Yannakoudakis,E.J. and Fawthrop, D.
An Intelligent spelling error corrector.- Information processing & Management.- Vol.19,No.2,Pp101-108,1983.