

التجارب الراهنة حول حوسبة النصوص التي تعتمد اللغة العربية

أ. موسى زموي
إطار في صندوق الإنماء الاقتصادي
والاجتماعي بالكويت

التجارب الراهنة حول حوسبة النصوص التي تعتمد اللغة العربية

موسى زمولي

الملخص

نطرح في هذا البحث - بعد التذكير بأهداف مشروع " حوسبة ذخيرة اللغة العربية " الذي يعتزم المجمع الجزائري للغة العربية إنجازه - إشكالية أهمية ومدى إمكانية إنجاز مثل هذا العمل الطموح، ونتطرق بعد ذلك إلى عرض وتقييم تجارب مماثلة تمت خلال العقدين الأخيرين، وكان من بين أهداف بعضها حيازة النصوص العربية في ذاكرة الحاسوب ونشرها عبر وسائط تكنولوجيا المعلومات، كما نسرد ما أنجز من أبحاث في مخابر فروع اللسانيات الحاسوبية خاصة منها التي تتعلق بموضوع معالجة اللغة العربية الطبيعية. ثم نتعرض من خلال هذا البحث إلى التطبيقات العملية المتمثلة في تجارب الشركة العالمية للإلكترونيات، والخطة المستقبلية للمنظمة العربية للتربية والثقافة والعلوم، ومشروع الوراق الإماراتي، وفعاليات القسم الثقافي بالمركز الجهوي للبرمجيات وهندسة المعلومات. كما نقارن هذه التجارب بتجارب خاصة باللغة الإنجليزية وأخرى متعلقة بمشروع "حوسبة اللغة الفرنسية" لمجمع اللغة الفرنسية.

ونخلص في تقييمنا لهذه التجارب إلى أن فاعليتها، بصفة عامة، لازالت دون حاجات المستخدم العربي.

وفي الفقرة الأخيرة نقترح بعض التوصيات حول طريقة إنجاز وتطوير مشروع حوسبة اللغة العربية الرائد للمجمع الجزائري للغة العربية.

إطلاقات أولية : ذخيرة اللغة العربية، حوسبة، توليد المعاجم حاسوبيا

1. مقدمة

إن التقدم السريع في تكنولوجيا المعلومات، بالإضافة إلى انتشار فكرة عولمة الثقافة، يملي على المؤسسات التي تعنى باللغة العربية أن تُعجّل في ربط جسر التواصل بينها وبين بقية العالم في هذا المجال.

لقد أنجزت أعمال قيمة في ميدان معالجة اللغة العربية الطبيعية، منذ السبعينات من طرف الباحث د. حاج صالح و فريقه (1) و خبراء آخرين، إلى جانب شركات تجارية ذات الاهتمام بالقضايا العربية[□].

- التأكيد على تبوأ اللغة العربية مرتبة هامة في صف اللغات المتطورة و ذلك بفضل مميزاتا المتمثلة خاصة في الاشتقاق والبنية المنطقية لقواعد الصرف والنحو و ثراء مفرداتها ومرونة التعامل معها مع قابليتها لاستقبال أول توليد المصطلحات الحديثة.

- إنجاز أبحاث معتبرة غايتها تحويل قواعد علم اللغة العربية إلى خوارزميات قابلة للحوسبة.

- توفر برمجيات عربية في الأسواق تضاهي نظيراتها في اللغات الأجنبية.

□. نذكر من بينهم د. نبيل علي (العلمية الكويت) ود. عبد القادر الفهري بجامعة محمد الخامس.

في هذا الإطار قرر المجمع الجزائري للغة العربية إنجاز مشروع "حوسبة الذخيرة اللغوية العربية" و يهدف المشروع حيازة أهم نتاج اللغة العربية من أدب وعلوم، على وسائط حاسوبية، لتوفير بنك معطيات نصية عربية محوسب يمكن نشره عبر شبكة الانترنت ومن خلال وسائل رقمية أخرى ليتسنى لأي مستخدم الاطلاع عليه بكل يسر.

سنعرض فيما يلي بعض الأمثلة حول التجارب الحالية ذات العلاقة بموضوعنا ومن بينها تطبيقات عملية لنماذج حاسوبية.

لنبدأ أولاً بتساؤلات حول جدوى و حظوظ إنجاز هذا المشروع الطموح.

2. الإشكالية

نرى أنه من الأجدر، في هذه المرحلة، أن نطرح ما يتراءى لنا من الأسئلة قبل الخوض في الأجوبة و اقتراح الحلول.

السؤال الأول : من أين نبدأ ؟

ج 1 : لقد ترك موضوع المصطلحات الحديثة، جانبا، ولقرون عدة، ليخضع لاجتهادات وإرتجالات المستخدم العربي دون أي سند، لينتج عن ذلك تراكم الأعمال وتأخر كبير في تحديث اللغة العربية ولا بد من مواجهة هذه الحقيقة قبل أو بالتوازي مع مشروع الحوسبة.

ولو طرح السؤال، حاليا، حول مفهوم عبارة اللغة العربية لكانت الردود مختلفة : هناك اللغة الفصحى المقررة في المدارس، وهناك ما يسمى "باللغة العربية العصرية" المستخدمة في مختلف وسائل الإعلام، ثم يأتي بعد ذلك دور اللهجات المتفرعة عنها.

السؤال الثاني : ما هي شروط انتقاء النصوص لضمها للذخيرة ؟ ومن هو

المستخدم المستهدف للذخيرة اللغة العربية المحسوبة ؟

ج2 : تعج الساحة الآن بمواقع عربية كثيرة على شبكة الإنترنت تجعل

بالإمكان الوصول إلى برمجيات عديدة مثل نظام القرآن الكريم بمختلف مزاياه ونظم المواريث والأحاديث النبوية الشريفة (الملحق 1). كما يمكن تصفح معظم الصحف العربية على الإنترنت في شكلها النصي ، أو الاطلاع على الشعر الجاهلي والمعلقات ، في إطار الوسائط المتعددة ، كما يمكن كذلك زيارة مواقع عربية مخصصة للأطفال وأخرى للمرأة ، وهناك مساحات على الشبكة للسياحة وغيرها (الملحق رقم 2). إلا أن المستخدم العادي للأنترنت في العالم العربي ما يزال ، في نظرنا ، في حاجة إلى معاجم عربية لتزويده بالمصطلحات العربية المرادفة لنظيرتها في الإنجليزية في ميادين التكنولوجيا والمالية ووسائل الإعلام وفي الطب وغيره ، كما هو بحاجة كذلك إلى موسوعات باللغة العربية في كل المجالات . باختصار فإن المستخدم العربي يطالب بعمل شمولي ومهني.

السؤال الثالث : ضرورة التنسيق.

ج3 : لقد أثبتت التجربة فشل محاولة التنسيق حتى بين مراكز البحث في بلد

عربي واحد بل حتى في نفس المؤسسة وذلك في معظم الحالات .

ويحضرني بهذه الخصوص قرارات ، أخذت منذ عدة سنوات في إطار الجامعة

العربية ، تتعلق بإنشاء شبكة بيانات حاسوبية عربية (2) ، مع تعيين لجنة متابعة

؛ ولحد الآن لا تزال حبرا على ورق وتنتظر من ينفذ عنها غبار النسيان !

السؤال الرابع : ما هي العوائق الفنية التي تعترض حيازة ذخيرة اللغة

العربية ؟

ج4 : حتى الآن وحسب علمنا توجد ثلاث طرائق للحصول على نص محوسب : إدخال كامل النص يدويا أو عن طريق المسح الضوئي للوثائق والتعرف عليها آليا، وأخيرا التعرف آليا على الكلام وتحويله إلى نص. ولا تزال هذه الطرائق الثلاث غير قادرة على الوفاء تماما بالغرض المنشود بسبب حجم ونوعية الوثائق المراد حيازتها. مما يستوجب الاستفادة من الوسائل الثلاث معا إلى أن يتوفر حل بديل مرضٍ □ .

3. أمثلة من التطبيقات الحالية

انطلقت الأبحاث في ميدان اللسانيات الحاسوبية في نهاية السبعينات وبداية الثمانينات. وكان كل الباحثين من مشرق العالم العربي إلى مغربه يعملون تقريبا على نفس المواضيع. وكان همهم الأول توحيد المحارف العربية، أو العمل على إنجاز منسق كلمات ومعالج للجداول وتوفير قاعدة بيانات تتعامل مع اللغة العربية للاستخدامات المكتبية وتطوير كل ذلك. وفي نهاية الثمانينات شرع الباحثون في تطوير خوارزميات لمعالجة اللغات الطبيعية.

وفي هذا المجال قدمت هيئات علمية مثل مركز البحث العلمي والتقني لتطوير اللغة العربية بالجزائر ومركز الحاسوب بجامعة محمد الخامس بالمغرب أعمالا نظرية

□. مثل نظام (Arabdox) الذي تقترحه شركة العالمية والذي ينطلق من المسح الضوئي للوثائق وفهرستها آليا مع إمكانية الربط بين الوثائق وإجراء بحوث موضوعية.

معتبرة. وحاولت تلك الهيئات وضع القواعد المناسبة بتطوير المحللات الصرفية والدلالية والمحللات النحوية التي تضبط اللغة العربية. وكمثال على ذلك يمكن الإشارة بالأبحاث المتعلقة بالمدرسة الخليلية الحديثة التي طورها د. حاج صالح وفريقه، كما سهر فريق المغرب والشركة العالمية وغيرهم على تحقيق تجربة التوليد الآلي للمعجم والتعرف آليا على المحارف وعلى الكلام المنطوق وكذا الترجمة الآلية، أما الشركات التجارية ومنها شركة العالمية الرائدة في هذا المجال وشركة (Alis) الكندية ومؤسسة ميكروسوف الأمريكية، فقد انحازت إلى الجانب العملي وأبلى بعضها فيه بلاء حسنا.

1.3. إنجازات شركة العالمية (3) و (4).

بدأت الشركة الكويتية "العالمية للالكترونيات" في الثمانينات بالاهتمام بالحاسوب العائلي (Home Computer) بتقديم برامج تثقيفية، تعليمية وترفيهية مثل برامج رحلة إلى مكة، تعلم الإنجليزية وكلمات ساحرة. وبداية من التسعينات حققت مؤسسة صخر (مجموعة مختصة بالبرمجيات تفرعت عن العالمية) نقلة نوعية بتوجهها نحو جيل جديد من البرمجيات يعتمد معالجة اللغات الطبيعية فطورت الشركة محلا صرفيا كان نواة لإنجاز برمجيات عربية مكتبية معتبرة (الملحق رقم 1)، على أقراص الليزر، مثل :

- برامج الناشر المكتبي والقرآن الكريم والحديث النبوي الشريف والمواريث
- المحلل النحوي
- المشكل الآلي
- المفهرس الآلي

- نظام الترجمة الآلية
- القارئ الآلي
- ونظام (Arabdox) الشامل لمعالجة الوثائق العربية مع فهرستها آليا وإمكانية الربط الموضوع كما أشرنا إلى ذلك سابقا.

وانتقلت الشركة في نهاية التسعينات إلى التركيز على شبكة الأنترنت كأكبر وسيلة متداولة في العالم. فأنشأت "السندباد" متصفح على موقع صخر (Sakhr.com) وخادوم عجيب (Ajeeb.com) الذي يسمح على سبيل المثال بتعريب بعض "البرتالز" وكان آخر منتجاتها الموقع الإعلامي المتميز (Johaina.com).

ولمجموعة صخر علاقات جيدة مع شركات تكنولوجيا المعلومات الدولية إذ تمثل الشركة العربية همزة الوصل بين العالم الغربي والمستخدم العربي. ولمزيد من المعلومات حول خدمات هذه الشركة يمكن الحصول على ذلك بتصفح المواقع التالية :

www.sakhr.com أو www.ajeeb.com أو www.johaina.com.

2.3. مشروع الوراق (www.alwaraq.com)

انتهى المجمع الثقافي بأبو ظبي من وضع الملامح الأساسية لمشروع الوراق وهو مشروع ثقافي رائد يستهدف إحياء التراث العربي من خلال نشر أبرز مقتنيات المكتبة العربية إلكترونيا على شبكة الأنترنت. وحسب تصريحات الشاعر نوري الجراح المشرف على المشروع فإن هذه الخطة تعتبر ثورة حقيقية بكل المقاييس، تهدف إلى تكوين أول موقع إستراتيجي من نوعه باللغة العربية على الأنترنت. فمن المعروف أن هناك أكثر من 2.5 مليار صفحة مكتوبة على الأنترنت في أكثر من 200 مليون موقع، هي عبارة عن

معارف ووثائق وأفكار تتصارع فيما بينها، والأغلبية منها إلى الآن باللغة الإنجليزية التي تحتل 86% مقابل 2.1% للغة الفرنسية، ويبقى 12% لباقي لغات العالم.

قسم المشروع إلى ثلاث مراحل : الأولى لحيازة كتب التراث، والثانية للمؤلفات الموضوعية بعد الحملة الفرنسية على مصر، وخصصت المرحلة الأخيرة للمؤلفات التي تم وضعها بعد الحرب العالمية الأولى حتى اليوم.

ويقول الجراح رئيس المشروع أنه قد تم إنجاز المرحلة الأولى نهائيا، واستطاع المشروع نقل الثقافة العربية من عالم الورق إلى أفق الضوء عبر شبكة الأنترنت، لأن من لا موقع له على الأنترنت لا موقع له على الأرض. ويضيف الجراح : إنه تم نشر 600 كتابا على الشبكة الدولية، بلغ عدد صفحاتها أكثر من مليون صفحة ومن المنتظر أن يبلغ عدد الصفحات نحو 10 ملايين صفحة.

وعن كيفية اختيار الكتب المرشحة للنشر على الشبكة قال مسؤول المشروع : نحن نحاول تغليب تراثنا العقلاني عن تلك التي خوض في الغيبيات وسوف تدخل الكتب في شكل نصوص لا صور ليسهل التعامل معها سواء بالحفظ أو الأرشفة، وفي المرحلة الحالية يتم إدخال المجالات التنويرية التي ظهرت في النصف الأول من القرن العشرين مثل : المنار والمقتطف والعصور والرسالة. وانتهى الجراح إلى القول بأن المشروع يسعى إلى تكوين جمعية عربية للنشر الإلكتروني تضم كل المواقع العربية للتنسيق بينها لحماية حقوق الملكية الفكرية.

ويمكن للباحثين الدخول على موقع الوراق لتصفح العديد من الكتب التراثية ومن بينها : الأصمعيات والإمتاع والمؤانسة، رسالة الغفران، نهج البلاغة، تاريخ الطبري، عجائب الآثار للجبريتي، رحلة البلاغة وغيرها.

3.3. مشروع المركز الجهوي للبرمجيات وهندسة المعلومات

(Ritsec.com.eg)

استفاد المركز من هبة قدرها 550000 د.ك. (1.650.000 دولار) قدمها له الصندوق العربي للإنماء الاقتصادي والاجتماعي في السنوات التالية 1993 ، 1994 و1997 لإنجاز نظام متكامل للمخطوطات العربية الإسلامية. في سنة 1995 ، صدر عن المركز مجموعة من الأقراص الضوئية بعنوان "ذخيرة الحضارة الإسلامية" (Treasure of Islamic civilization) وهي مجموعة صور لمخطوطات عربية محفوظة بجامعة برنستون الأمريكية. وفي سنة 1997 صدر عن المركز مجموعة أخرى من الأقراص الضوئية حول عينة من الصور لمخطوطات دار الكتب المصرية.

4.3. مشروع مركز مكي جمعة الماجد (الإمارات العربية المتحدة)

يعمل هذا المركز منذ سنوات على إحياء التراث العربي وهو أيضا مركز توثيق ومعلومات. و من إنجازاته على سبيل الخصوص، تصوير بوساطة كاميرا رقمية، أكثر من 3000 مخطوط تابع للمكتبة الوطنية الجزائرية في سنة 2000.

5.3. آفاق عمل المنظمة العربية للتربية والثقافة والعلوم (2003.2010)

في 16 يونيو (جوان) 2001، استضافت الكويت ملتقى خبراء للبحث في الخطة المستقبلية لعمل المنظمة وجاء الحاضرون بتوصيات أهم ماله صلة بموضوعنا ما يلي :

- العناية باللغة العربية وتوظيف أساليب واستراتيجيات جديدة في تدريسها، مع الاهتمام بنشر القيم العربية واللغة العربية في الدول الإسلامية من خلال تعليم اللغة العربية لغير الناطقين بها،
- المساهمة في إقامة البنية التحتية لصناعة ثقافة عربية،
- إحياء التراث العربي بمعناه المادي والمعنوي تقديمًا وصياغة و ترميماً،
- دفع جهود التنظير اللغوي،
- توثيق التراث اللغوي والاستخدامات الحديثة للغة العربية،
- تطوير أساليب بناء المعاجم،
- العناية بالترجمة من اللغات الأجنبية إلى اللغة العربية،
- التأكيد على أن استخدام التقنيات الحديثة أصبح ضرورة ملحة ومدخلا رئيسا للاندماج في العصر،
- الاستفادة من شبكة الانترنت في التواصل بين الدول العربية ومختلف مؤسسات البحث العلمي، ومراكز الدراسات الاستراتيجية في الدول العربية،
- تحديث منهجيات تعلم وتعليم اللغة العربية،
- إقامة قاعدة ببلوغرافية للتنظير اللساني الحديث،
- تعميم النظم البرمجية لمعالجة اللغة العربية آليا،
- وضع القوانين اللازمة لحماية الملكية الفكرية.

6.3. شبكة بيانات المؤتمر الإسلامي (OICS – Net) [□]

يعد مشروع منظمة المؤتمر الإسلامي من الشبكات الحاسوبية الرائدة في العالم الإسلامي حيث انطلق المشروع في منتصف الثمانينات باستخدام بنية الاتصالات التابعة لشركة (SITA) [□]. وبدأت المنظمة بوضع محطات (Help desks) لشبكتها في بعض الدول الإسلامية والتنسيق مع البلدان الأخرى لتبادل المعلومات التي تهتم العالم الإسلامي. وبعد ظهور شبكة الانترنت سعت المنظمة لإعادة تصميم شبكتها للاستفادة من البنية الجديدة لتقانة المعلومات ولجعل المعلومات المتعلقة بالعالم الإسلامي في متناول الجميع.

7.3. مركز التعريب التابع للجامعة العربية

يهتم المركز، الذي يوجد مقره الرباط، منذ سنوات بدراسة وتطوير المصطلحات العربية خاصة العلمية والتقنية منها. كما قام المركز بإصدار معاجم متخصصة. ومن المفترض أن يكون للمركز، الذي يمارس مهامه منذ حوالي 20 سنة، بنك بيانات بالمصطلحات العربية والأجنبية يمكن للمجمع الجزائري للغة العربية الفتي والمجلس الأعلى للغة العربية الاستفادة منه.

8.3. تجربة الكاتب الطاهر وطار (www.aljahidhiya.ass.dz)

قام الروائي الجزائري الطاهر وطار بتعلم لغة (Java Script)، بمفرده لإنجاز موقع للمؤسسة التي يرأسها. ونشر عبر صفحات هذا الموقع مجلة الجاحظية وأخبار الأدب والفنون بالإضافة إلى رواياته الشهيرة. وحسب علمنا فإن سي الطاهر روائي قدير ولكن لم يكن على دراية بتكنولوجيا المعلومات وحقق بمفرده هذا الإنجاز مما يعني أن ذلك سيشجع الكتاب والأدباء للاحتذاء به خاصة بعد قيام ميكروسفت

□. Société Internationale de Télécommunications Aéronautiques.

بتحديث برنامج (Front Page)، في الإصدار الأخير : (Ms office XP) حيث وفرت قوالب لإنشاء مواقع على الانترنت بسهولة ويسر. كما لا يفوتنا هنا أن نشيد بدور مركز [□] (CERIST) الذي استضاف موقع الجاحظية.

9.3. الصندوق العربي و مشروع " المعجم العربي الحديث "

في سنة 1987، خصص الصندوق العربي للإنماء الاقتصادي والاجتماعي، 100 ألف دينار كويتي (حوالي 300 ألف دولار) كقسط أول لإنجاز معجم عربي حديث واستقطب الصندوق خبراء وأعضاء في مجامع اللغة العربية، تحت إشراف د. عثمان سعدي (عضو مجلس إدارة الصندوق العربي آنذاك).

وقامت للجنة بتحديد أهداف المشروع ومراحل تنفيذه كما تم ضبط منهجية العمل. وكان المشروع يهدف في النهاية إلى إنشاء مؤسسة دائمة مستقلة، على غرار مؤسسة (لاروس)، لتحديث المعاجم العربية دوريا. وكان مقررا أن تتم حيازة البيانات على الحاسوب. وتم بالفعل إنجاز جزء من المشروع حسب المنهج العلمي الذي اعتمده اللجنة. إلا أن العمل توقف بسبب حرب الخليج الثانية. من جهة أخرى، وعلى سبيل المقارنة، نعرض في ما يلي تجربتين تخصان اللغتين الإنجليزية والفرنسية :

10.3. معجم (Colins Cobuild English Dictionary)

وهذا المعجم من إصدارات دار جامعة كولنز - برمنغهام، التي تتوفر بها قاعدة بيانات لغوية. وقد تم توليده حاسوبيا من هذه القاعدة. ومن أهداف هذا المعجم

□. Centre des Etudes et de Recherche en Information Scientifique et Technique.

مساعدة الطلبة والأساتذة في أبحاثهم وأعمالهم، ويعد ثمرة جهد فريق من خبراء جامعة برمنغهام الذي كرس سبع سنوات من العمل الدؤوب، بالاستعانة بكل خدمات الحاسوب لتحليل اللغة الإنجليزية المتداولة حالياً مستخدمين في ذلك ملايين النصوص ومختلف نماذج استخدام اللغة الإنجليزية.

11.3. المشروع الفرنسي "حوسبة ذخيرة اللغة الفرنسية"

(Trésor de la langue Française Informatisé)

ارتأينا أنه من اللائق عرض هذه التجربة للاطلاع على منهجية هذا العمل الذي تتقارب أهدافه مع مشروع المجمع الجزائري الفتى.

وللاطلاع على مراحل وأماكن إنجازه وعلى الصعوبات التي عانى منها اللسانيون والمؤسسات ذات العلاقة بالمشروع، يمكن تصفح الموقع التالي :

www.ciril.fr/~masina/TLF ، وكذا الاطلاع على الملحق رقم 3.

4. تقييم هذه التجارب

بالاطلاع على قائمة البرمجيات (الملحق 1)، وباستثناء نظام القرآن الكريم والحديث الشريف، التي تمت الموافقة على نشره بعد فحص دقيق من طرف الجهات الإسلامية المعنية بالموضوع، فإن معظم باقي البرمجيات يمكن مناقشة محتواها ونوعيتها. وعلى سبيل المثال فإن نظام التعرف آلياً على المحارف العربية (OCR) لا زالت نسبة جودته غير كافية للاستخدام المهني. وإذا كان بالإمكان تحسين منتج الـ OCR عن طريق البرمجة، فإن قضية تحسين الترجمة الآلية، مثلاً، تكمن في التحليل الضمني للكلمة المراد ترجمتها، وهذه قضية شائكة من قضايا

التحليل الدلالي. فعلى سبيل المثال نجد كلمة "علم" التي تشير إلى عدة معانٍ من علم إلى علم. ولمعرفة معناها الدقيق في النص فليس لنا سوى اللجوء إلى التحليل الضمني أو تشكيل الكلمة. ويبدو أن شركة العالمية نجحت في إنجاز برمجيات لتشكيل النصوص العربية آلياً.

أما بخصوص المواقع العربية (الملحق رقم 2)، فمعظمها لا يلبي الغرض المنشود لا من حيث شمولية المواضيع ولا من حيث نوعيتها وأغلبها يعالج مواضيع إسلامية تراثية بشكل سطحي أو تجاري إعلاني أو ترفيهي. أما مجموعة صخر التي أنجزت ثلاثة نظم : (ajeeb.com، Johaina.com و www.al-islam.com) فإنها توفر أدوات جيدة للهواة والمحترفين في آن واحد وتستحق الزيارة والتقييم. كما لنا وقفة بخصوص مشروع الوراق الطموح وهو أقرب عمل إلى مشروع المجمع الجزائري وللمشروع بداية حميدة لكن من الصعب الحكم له أو عليه حالياً.

وبخصوص مراكز اللسانيات ودور البحث النظري فلن نجدوا فرصة أحسن من مشروع حوسبة ذخيرة اللغة العربية لتطبيق نظرياتهم وخدمة مستخدمي اللغة العربية.

5. التوصيات

مما لاشك فيه أن الهوة بين مجامع اللغة العربية وحاجة مستخدمي هذه اللغة، كبيرة وقائمة منذ قرون. ولذا يتحتم علينا إعادة النظر في المشاريع ذات العلاقة بموضوع توفير الذخيرة العربية على الانترنت ومنهجية البلوغ لهذه الغاية. ومن استقرائنا للتجارب السابقة والمتعلقة بالمبادرات والنوايا الحسنة لخدمة القارئ العربي خلصنا إلى أن هذه المحاولات لحوسبة التراث العربي، على سبيل الخصوص، لا زالت جنينية في بدايتها. وتبقى فردية دون أي تنسيق بين مختلف

الجهات مما يفقدها شيء من المصادقية مع مصرة عدم الاستفادة من أخطاء الآخرين. في هذا الإطار وبخصوص خطة تنفيذ مشروع "حوسبة ذخيرة اللغة العربية" نقترح ما يلي :

أولاً : الجانب الأكاديمي

- اختيار المؤلفات : يجب أن يتم ذلك بواسطة فريق متخصص يشمل كل فروع المعرفة لمراجعة الحصيصة الموثقة للتراث العربي الإسلامي.

- اختيار المنهجية : حيث أن مصادر الذخيرة العربية مختلفة كما ونوعاً ولأن الاختيار المتأني السليم في البداية سيعود على الجميع بالفائدة المؤكدة ولا يمكن استعمال طريقة الخطأ والتجريب في مثل هذا العمل الضخم، لذا يتحتم على أصحاب القرار، ومنذ البداية اختيار المنهجية اللائقة من أجل احازة الوثائق والأدوات المعلوماتية التي سترافق المشروع طول حياته والعمل على احترام هذه المنهجية مع احتمال مراجعتها، بالمشورة، إذا اقتضى الأمر لتتماشى مع التطور التكنولوجي.

- تحديد الأولويات : بما أن لكل مشروع قيود مادية أو زمنية فلا بد إذا من تحديد الأولويات وتخطيط مراحل تنفيذها. على سبيل المثال يمكن في مرحلة متوسطة المدى من المشروع العمل على حيازة الذخيرة واستخلاص مكنز يتم من خلاله التوليد الآلي لمعجم عربي حديث ينشر على شبكة الانترنت.

- الاستناد على أبحاث اللسانيات : إن أهم ما تميز به عصرنا، بعد تكنولوجيا المعلومات، هو تطور فرع اللسانيات ليصبح علماً مستقلاً من علوم المعرفة وإن النماذج

الرياضية التي وصلت إليها مراكز البحث كما أشرنا إليه سابقا لا يمكن إلا أن تخدم قضيتنا هذه وأن تضمن مصداقيتها. ولذا يجب أن تكون هذه النماذج الرياضية هي النواة التي يحيط بها باقي النسيج.

ثانيا : الجانب الفني

على خبراء تكنولوجيا المعلومات اختيار، قدر المستطاع، أحدث أدوات حاسوبية تم اختبارها وأكثرها تطابقا وتمشيا مع نظم التشغيل الحالية والمنظورة. أما ميكانزمات حيازة الذخيرة فيجب أن تكون مجزئة (Modular) ومرنة بحيث يمكن تحديث جزء من العملية دون المساس بالمشروع ككل، أو إضافة حصيلة قد تأتي من طرف آخر أو تجربة سابقة. كما يجب أن يكون هناك منذ البداية تصور للمنتج النهائي حتى من حيث نوع وسائط نشره و توزيعه (ليس فقط على الانترنت بل لا بد و أن تراعي احتياجات المعوقين، مثلا).

ثالثا : أعمال التشغيل والصيانة

تحديث النظام وصيانتته ليس بالأمر الهين مع تقدم أبحاث اللسانيات وتطور التكنولوجيا من جهة ومع توفر كم هائل من المعلومات التراكمية يوميا. ولضمان استمرارية المشروع وتجانس وحداته لا بد من تخصيص جهد كبير لهذا العمل كما يجب للمحافظة على سلامة المشروع اعتماد قواعد وقائية وأمنية مؤكدة تضمن استمرارية المشروع في كل مراحله.

رابعا : الموارد البشرية

لا يمكن أن يضمن نجاح مثل هذا العمل المضني والجاد إلا عن طريق فريق من مختلف فروع المعرفة مؤمن بأهمية هذه القضية ومتفرغ لها. وعلى أصحاب القرار توفير الجانب المادي لهذا الفريق وظروف عمل تشجع على خلق روح الإبداع والإنتاجية. وينطبق ذلك على كل المستويات سواء تعلق الأمر بخبراء أم بعمال بسطاء.

خامسا : الجانب المالي

لقد بات معروفا لدى الكثير بأن نسبة نجاح المشاريع الطموحة وذات المدى المتوسط أو البعيد، في محيطنا، ضعيفة وذلك بسبب الموارد المالية أو بسبب سوء استخدامها. ويجب تظافر الجهود لإنجاز المراحل الأولى من المشروع الذي قد تأتي بعوائد تسمح بالاستقلالية المالية للمشروع. علما بأن الغرض المنشود لا يمكن أن يكون تجاريا.

سادسا : الجانب التشريعي

على الجهات المعنية بالموضوع أن تعمل بالتكامل على سن القوانين التي تحث المثقفين على توفير إنتاجهم على وسائل رقمية لتسهيل مهمة الاستفادة منها. وأخيرا يبقى الهدف المنشود هو إنجاز نظام يضمن إستمراريته وتطوره مستقبلا.

المراجع

1. Al Hadj Saleh, A, . المدرسة الخليلية الحديثة ومشاكل علاج العربية بالحاسوب ، 1989, Proceedings of the International Conference on « Arabic computational linguistics, held in Kuwait Scientific Research Center, Kuwait, 1989.
2. Zemouli, M, 2000, Information technology in Arab Countries : Assessment and perspectives. International conference on “ Arab Development Challenges of the new Millennium ”, organized by the Arab Planning Institute (API) in Kuwait and the University of Mohammed V, June 26-28, 2000 in Rabat, Morocco.
3. Nabil, A, 1988, The Arabic Language and the Computer (اللغة العربية) (والحاسوب), Taarib.
4. Abu al-Azm, A. Arabic Language and Automatic Processing, ‘ اللغة العربية ’, « و المعالجة الآلية برنامج صخر نموذجاً Fikr wa naqd, n° 31, September 2000.

الملاحق

الملاحق 1 : قائمة ببعض البرامج النصية العربية المحسوبة

ملاحظات	الجهة المصدرة للمنتج	تسمية المنتج
عمل قيم يقدم النص الكامل للقرآن مع تفسيره وترجمته بعدة لغات وطريقة لتحفيظه.	شركة العالمية	القرآن الكريم
برامج شامل ودقيق لأحكام المواريث	//	المواريث
برنامج كامل للأحاديث النبوية الصحيحة	//	الحديث
إنجليزي عربي وناطق	//	القاموس
مليون كلمة، عربي - إنجليزي	ATA (الكويت)	الوافي
قاموس ابن منظور	جمعية إحياء التراث الإسلامي (الكويت) 1999.	لسان العرب
يقوم بإلقائها ممثلون مشهورون	//	المعلقات السبع
مجموع من مؤلفات ابن القيم	مكتبة ابن قيم الجوزي، الكويت 1999	أعمال ابن قيم الجوزي
في إطار الحوسبة	وزارة الأوقاف الكويت	الموسوعة الفقهية

//	وزارة الاعلام الكويتية	تاج العروس
//	مؤسسة عبد العزيز البابطين - الكويت	موسوعة الشعر العربي

الملحق 2: عناوين بعض المواقع العربية على الانترنت

Portal Name	Portal Address
Algerian General Information Database	www.algeriainfo.com
One hundred Arab sites	www.arank.com/
Taher outar	www.aljahidhiya.ass.dz
Naseej	www.naseej.com
Ayna	www.ayna.com
4Arabs	www.4arabs.com
Arabseek	www.arabseek.net
Konouz	www.konouz.com
Gulf Palace	www.gulf-palace.com
Al Saha	www.alsaha.com
Planetarabia	www.planetarabia.com
Moheet	www.moheet.com
Saudi Online	www.saudionline.com.sa
Middle East Inernet Pages	www.middle-east-pages.com
Arabia Online	www.arabia.com
1001sites	www.1001sites.com
Arab Net	www.arab.net
Wahat	www.homestead.com
Arab View	www.arabview.net
Maktoob	www.maktoob.com
Horizons	www.ihorizons.com
Dabbagh Information Technology	www.dit.net
Arab Business	www.arab-business.com
Lexic Orient Front Page	www.i-cias.com
Arabvista	www.arabvista.com
Al Jareh	www.aljareh.net
Arabic 2000	Arabic2000.com
Sites for all	www.sitesforall.com
Arabist	www.arabist.com
All Saudi	www.all-saudi.com
Ya Salam	yasalam.com
Bawaba	bawaba.com
Zorona	www.zorona.com

Nawal	www.nawal.com
Nana	www.nana.co.il
Arabji	www.arabji.com
Sinsal	www.sinsal.com
Arabcafe	www.arabcafe.com
Zooloo	Zooloo.co.il/Arabic
Daleel	Adleel.com
Kufta	www.kufta.com
Arabic Seek	www.seek.com
Fares	www.fares.net
Saudi Land	www.saudiland.com
Yalla	www.yalla.com
Raddadi	www.raddadi.com
Al Aseel	www.alaseel.net
Al Multaga	www.almultaga.com
Arabichat	www.arabichat.net
Arabian Oasis	www.arabianoasis.net
Kunoozy	www.kunoozy.com
Cafe Arabia	www.cafearabia.com
Kuwaitvoice	www.kuwaitvoice.com
Sandroses	www.sandroses.com
Quality Net	www.qualitynet.net
Jeeran	www.jeeran.com
Inter Nec	www.inter-nec.com
Mauje	www.mauje.com
Bayt	www.bayt.com
Khaleej	www.khaleej.com
Khayma	www.khayma.com
Yafea	www.yafea.com
Saudi Links	www.saudilinks.com
Eqla3	www.eqla3.com
Awalnet	www.awalnet.com
Lakianti	www.lakianti.com
Jahra Net	www.jahra.net
Arabna	www.arabna.com
Beseen	www.beseen.com
Les enchères au Maroc	www.dlala.com
Arabia online Guernsey Ltd.	www.arabia.com
Al Alamiah	ajeeb.com
Al Alamiah	paintings.ajeeb.com
Arabic news	www.arabicnews.com/
Lybia	www.hmnet.com/africa/libya.html
Al Alahmiah	tarjim.ajeeb.com/ajeeb/
Al-Alamiah	johaina.ajeeb.com
Arabic Media Internet Network	www.amin.org-jourmag/
African links	www.afrik.com/

Egyptian children web	www.horus.org.eg
Inter-Arab Investment Guarantee corporation	www.iaigc.org
Qatar, Ministry of Awqaf and Islamic Affairs	www.Islamweb.net
Afkar Information Technology Ltd.	www.Islamonline.net
Kuwait Studies & Reseach Center	www.crk.org
Algerian newspaper	www.comfm.fr/presse/dz.html
Harf	www.al-islam.com

Annex 3 : French Automated Dictionary

Trésor de la langue Française informatisée, T.XIV, interrogeable sur internet à www.ciril.fr/~mastina/TLF.

Aspects de la base informatisée du Dictionnaire de l'Académie française de 1694 Russon Wooldrige, University of Toronto.

© R. Wooldrige, juin 1998

[Version révisée d'une conférence présentée à Paris, le 15 mai 1998, au Groupe d'études en Histoire de la langue française.]

1. Modalités de saisie
 - 1.1. Compétence du claviste
 - 1.2. Explicitation des co-vedettes elliptiques
2. Aspects de la genèse du dictionnaire
 - 2.1. Ordre alphabétique
 - 2.2. Renvoi et article
 - 2.3. Oublis rattrapés
 - 2.4. Regroupements étymologiques
3. Nomenclature dictionnaire et lexique réel
 4. Nomenclature cachée et mots oubliés
5. Adéquation typographie / microstructure
6. Le texte comme dictionnaire et corpus
7. Conclusion

الملحق رقم 4 : تجربة برمجيات الترجمة

النص المراد ترجمته من الانجليزية إلى اللغتين العربية والفرنسية

Subject : CEO OF HEWLETT PACKARD one Islam !!!

- Dear friends,
- I have received this email from a friend and I am passing it to you.
- This is the best article I read after September 11 tragic events.
- When you read this article you know why Hewlett Packard, the great global
- Techology company, is putting has lady, this lady in the world of men, have its
- CEO
- What has great business leader and thinker she is, and what has harmony between
- Business and ethics she has.
- You will enjoy reading the full article if you like reading, morale, ethics,
- Leadership, business, religion, and humanity.
- Moussa Zemouli
- Carly Fiorina, the CEO OF HEWLETT PACKARD , recently mountain stream has speech (september 26, 2001) defining the relevance of leadership in today world ("WHAT DOES FUTURE OUR DEMAND OF LEADERS TODAY ?"). The example she used was Islam between 800 and 1600 It. This gives a look at Islam which is quite different from the ones we have seen /heard/ read in the media since September 11. The Raising (finding) part of her speech is quoted below. For the full speech, please go to : <http://www.hp.com/hpinfo/execteam/speeches/fiorina/minnesota01.htm>

of here the final leaves of Fiorina speech :

- Have business leaders, have we are faced with questions of life and death
 - Rather than how much our stock is worth, the significance of our business
- Contribution to the world may be increased. And that is has good thing.
 - I ll end by telling a story.
- There was ounce a civilization that was the greatest in the world.
 - It was able to create has continental high-octane gasoline-state that stretched from ocean to ocean, and from northern climes to tropics and deserts. Within its Dominion.
- Lived hundreds of millions of people, of different creeds and Origins ethnic.

ترجمة النص بواسطة Ajeeb.com

الموضوع : سيو هوليت باكارد الإسلام واحد !!!

- الأصدقاء الأعزّاء
- قد تلقيت هذا البريد الإلكترونيّ من صديق و أناوله لك.
- هذه هي أفضل قطعة قرأتها بعد 11 سبتمبر أحداث مأساويّة.
- عندما تقرأ هذا المقال أنت تعرف لماذا هوليت باكارد، العظيم عالميّ.
- شركة التكنولوجيا، يضرب لدى سيّدة، هذه السيّدة في عالم الرجال، قدها.
- سيو.
- ما لدى قائد تجارة عظيم و مفكّر هي، و ما لدى ائتلاف بينهما.
- التجارة و الأخلاقيات هي قد

- سوف تستمتع بقراءة المقال الكامل إذا أحببت القراءة، الروح المعنوية، الأخلاقيات.

- القيادة، التجارة، الدين و البشرية.

- موسى زمولي

- كارلي فيورينا، سيو هوليت باكارد، حديثا لدى جدول الجبل خطبة (26 سبتمبر، 2001) تعريف صلة القيادة بالداخل اليوم عالم (ماذا مستقبل ملكنا يطلب من القادة اليوم ؟). كان المثال الذي استعملته الإسلام بين 800 و

على

1600 آي تي. هذا يعطي نظرة عن الإسلام المختلف إلى حد ما عن الذين قد رأينا/ سمع/ قرأ في الإعلام منذ 11 سبتمبر. الرفع (تحقيقات) جزء خطبتها يستشهد إلى أسفل. للخطبة الممتلئة، من فضلك اذهب إلى : <http://www.hp.com/hexecteam/speeches/fiorina/minnesota01.htm> . من هنا

الورق النهائي من خطبة فيورينا :

- لدى قادة تجارة، قد نواجه نحن مع / ب أسئلة الحياة و الموت

- بدلا من كم سهما مستحق، مغزى تجارتنا

- قد تكون المساهمة إلى العالم.

ترجمة نفس الفقرة بوساطة برمجيات : Reverso Pro Softissimo

Sujet : PDG D'HEWLETT PACKARD un Islam ! ! !

- Chers amis

- J'ai reçu cet email d'un ami et je vous laisse le passe.

- C'est le meilleur article 1 lu après le 11 septembre ddes évènements tragiques.
- Quand vous lisez cet article vous savez pourquoi Hewlett-Packard, le grand global { *mondail* }.
- La société de technologie, met a la dame, cette dame dans le monde d'hommes, avoir son
 - PDG
- Ce qui a le grand leader économique { * chef d'entreprise* } et le penseur elle est et ce qui a l'harmonie entre
 - L'Affaire et l'éthique elle a.
- Vous aimerez lire le plein article si vous aimez lire, la morale, l'éthique,
 - Direction, affaire, religion et humanité
 - Moussa Zemouli

De voirement Fiorina, le PDG d'HEWLETT PACKARD , récemment le flot de montagne a le discours (le 26 septembre 2001) la définition de la pertinence de direction dans l'aujourd'hui monde (" QU'EST-CE QUI FAIT L'AVENIR NOTRE DEMANDE DE LEADERS AUJOURD'HUI ? "). L'exemple qu'elle a employé était l'Islam entre 800 et sur 1600 CELA. Cela jette un regard à l'Islam qui diffère tout à fait de ceux nous avons vu / a entendu/ lu dans les médias depuis le 11 septembre. Le levage (la découverte) la partie de son discours est cité ci-dessous . Pour le plein discours, allez s'il vous plait :

<http://www.hp.com/hpinfo/execteam/speeches/fiorinaminesota01.htm>

D'ici les feuilles { *congrés* } finales de discours Fiorina:

- " Ayez des leaders économiques { *chefs d'entreprise* }, faites faites face nous sommes avec les questions de vie et la mort.
- Plutôt que combien de notre réserve { *valeur* } vaut, la signification de notre affaire.
- La contribution au monde peut être augmentée. Et c'est a la bonne chose.

- Je II fin en disant une histoire.
- Il y avait l'once une civilisation qui était la plus grande dans le monde.
- Il était capable de créer a l'essence d'octane haut continentale-déclarent que s'est étiré de l'océan à l'océan et de climats du nord aux tropiques et des déserts. Dans son Dominion.
- Vécu centaines de millions des gens, de credos différents et ethnique Origines.

