

## المحتوى العربي على العنكبوتية العالمية؛ منظور معلوماتي

أيمن شعبان الدكروري

باحث بالماجستير، جامعة القاهرة

[ayman\\_shaban@hotmail.com](mailto:ayman_shaban@hotmail.com)

### مستخلص

دراسة حول المحتوى الرقمي العربي على الإنترنت من وجهة نظر علم المعلومات، تتناول الدراسة اللغة العربية كأساس للمحتوى الرقمي العربي على الإنترنت، حيث تعرض للسمات العامة للنصوص العربية على العنكبوتية العالمية وخصائص اللغة العربية على الإنترنت والضوابط التي ينبغي الإلتزام بها عند النشر الإلكتروني وقواعد التشكيل، والنظام العددي في اللغة العربية، وتناولت الدراسة أيضا المحتوى العربي على الإنترنت وفئاته وأنواع الوثائق المتاحة على الإنترنت.

### الاستشهاد المرجعي

أيمن شعبان الدكروري. المحتوى العربي على العنكبوتية العالمية؛ منظور معلوماتي . -  
Cybrarians Journal . ع 23، سبتمبر 2010 . - تاريخ الإطلاع <أكتب هنا تاريخ  
الإطلاع على البحث> . - متاح في: <أنسخ هنا رابط الصفحة الحالية>

### 1. تمهيد

هناك حاجة ماسة إلى منظور عربي للمحتوى على العنكبوتية العالمية، بوصفه أحد اهتمامات علم المعلومات، وبما يسهم في الوقوف على الوضع الراهن لطبيعة وخصائص هذا المحتوى، والتعامل معه حفظاً وبحثاً واسترجاعاً. وفي هذا الصدد يتم

تسليط الضوء على مجالات المحتوى العربي على العنكبوتية العالمية، وفئاته، وأنواعه، والسمات العامة التي تميزه عن غيره من المحتويات الأخرى. ولما كانت اللغة العربية ذات خصائص بنائية ونحوية وإملائية قد تتشابه مع (أو تختلف عن) غيرها من اللغات، فقد أصبح لزاماً التطرق لهذه الخصائص، وتسليط الضوء عليها، بحيث تكون نقطة الانطلاق عند تناول المحتوى العربي على العنكبوتية العالمية.

## 2. مجالات المحتوى على العنكبوتية العالمية

يمكن تصنيف مجالات المحتوى العربي على العنكبوتية العالمية في ثلاثة محاور رئيسية:

أ. الأعمال والتجارة بين المؤسسات وبعضها البعض، وبين المؤسسات والأفراد، وبين الأفراد والأفراد.

ب. الخدمات العامة: من إدارة عامة إلكترونية، وتعليم إلكتروني، وصحة إلكترونية، وثقافة إلكترونية، ونشاط اجتماعي، يعتمد على الحوار وتبادل الآراء عبر وسائط تلك الشبكة.

ج. الإعلام والترفيه في الإنتاج التلفزيوني، والترفيه السينمائي، والموسيقى، والمجلات، والصحف، والكتب العامة والتعليمية، والألعاب، والتسلية.<sup>(1)</sup>

## 3. فئات المحتوى على العنكبوتية العالمية

تتفاوت فئات المحتوى العربي؛ فهناك المحتوى المتفرق fragmented الذي يتم تجميعه من مصادر متناثرة ومتنوعة ومتفاوتة في درجة معالجته، من البيانات الخام إلى المعلومات المعالجة بالفعل. وهناك المحتوى الخام ذي الطابع السردى الخطي narrative. إضافة إلى المحتوى سريع الزوال الذي تقوم عليه الصحف ووسائل الإعلام الجماهيري، في مقابل المحتوى الأكثر ثباتاً الذي تنتجه الدوريات العلمية.

كما يتراوح هذا المحتوى بين معلومات صريحة ومعلومات ضمنية تفهم من السياق ومعلومات تضاف إلى هذا المحتوى من وقت لآخر، وقد تكون هذه المعلومات صادرة عن شخص أو حكومة أو مؤسسة معينة. وأيضاً قد يُقدّم المحتوى للعامة دون تمييز، كما قد يقدم لأشخاص بعينهم، سواء في صورته الخام أو بصورة شبه معالجة أو بمعالجة كاملة، أو قد يتخذ هذا المحتوى الشكل النصي أو العددي أو الرسوم البيانية أو

الرسوم التخطيطية أو الخرائط أو الصور الثابتة أو المجسمات أو الصور المتحركة أو الفيديو أو الصور التشكيلية. ومن الممكن أن يصدر المحتوى بشكل سري أو قد يصدر للكافة أو قد يكون محدوداً في توزيعه. ويوضح الجدول رقم 1 هذه الفئات.

الجدول رقم 1 فئات المعلومات التي يتضمنها محتوى العنكبوتية العالمية<sup>(2)</sup>

من حيث طبيعتها	ب. معلومات ثابتة	من حيث النسق الرمزي
أ. معلومات حكوماتية	ج. معلومات ديناميكية	أ. معلومات رقمية
ب. معلومات مؤسساتية	من حيث علاقتها بالنص	ب. نصوص
ج. معلومات مهنية	أ. معلومات صريحة explicit	ج. رسوم بيانية
د. معلومات شخصية	ب. معلومات ضمنية implicit	د. رسوم تخطيطية
من حيث نوعية المستفيدين	ج. معلومات مضافة added	هـ. خرائط
أ. معلومات للعامّة	من حيث درجة السرية	و. صور ثابتة ومجسمة
ب. معلومات للمتخصصين	أ. معلومات للكافة	ز. صور متحركة و فيديو
ج. معلومات للنخبة	ب. معلومات محدودة التوزيع	ح. صور تشكيلية
من حيث مستوياتها	ج. معلومات سرية	ط. كلام وأصوات وموسيقى
أ. معلومات تشغيل	من حيث درجة المعالجة	من حيث مدى تماسكها وتعاقبها
ب. معلومات إدارة (متوسطة وعليا)	أ. معلومات غير معالجة	أ. خطية سردية
من حيث معدل تغيرها	ب. معلومات شبه معالجة	ب. غير خطية nonlinear
أ. معلومات متطايرة volatile	ج. معلومات معالجة (سابقة التجهيز)	ج. متفرقة

وتجدر الإشارة إلى أن هناك خصائص لوثائق العنكبوتية العالمية أحياناً ما تسبب مشكلات في التصفح والاسترجاع، وهي:

أ. أن وثائق العنكبوتية العالمية تتسم بالحد الأدنى من التنظيم، وعادة ما تكون الوثيقة الواحدة مكتظة بالمعلومات.

- ب. احتمالية اشتغال الموقع الواحد على محتوى بصيغ أو تراكيب متعددة.
- ج. قلما يكون هناك أمر واحد يستخدم لتفريغ جميع المعلومات، وربما يتعين على المستفيد أن يتصفح كل موقع على حدة.
- د. إمكانية اختلاف المخرجات من نظام تصفح لآخر، ومن طريقة تفريغ لأخرى.<sup>(3)</sup>

#### 4. أنواع الوثائق على العنكبوتية العالمية

هناك ثلاثة أنواع من وثائق العنكبوتية العالمية:

##### أولاً: الوثائق الثابتة Static Documents:

- أ. تستخدم لغة تهيئة النصوص الفائقة HTML في عرضها.
- ب. تقوم بعرض المعلومات فقط دون أي تدخل من المستفيد.
- ج. يتم تحريرها بواسطة المسئول عن الوثيقة.
- د. تبقى كما هي دون تغيير إلى أن يتم تحريرها أو تعديلها أو حذفها أو الإضافة إليها.
- هـ. تتميز بسرعة إنشائها وانخفاض تكلفة استضافتها.
- و. من عيوبها أنها تتطلب خبرة تقنية لتحديث محتوياتها.
- ز. عادة ما يطلق عليها العنكبوتية السطحية surface web.

##### ثانياً: الوثائق الديناميكية Dynamic Documents:

- أ. يتم إنتاجها بواسطة نادل مراسم نقل النصوص الفائقة HTTP server.
- ب. تستعين بأي من لغات البرمجة: PHP، أو ASP، أو CMF ... لعرض المخرجات أو المحتوى للمستفيد.
- ج. تتميز بتفاعلها مع المستفيد بحيث يكون هناك نوع من التشارك بين القائمين على المحتوى وبين المستفيدين منه.
- د. تتسم بسهولة تحديثها وإضافة محتوى جديد لها، ما من شأنه اجتذاب عدد أكبر من المستفيدين وكثافة زيارات محرركات البحث لها.

٥. من عيوبها بطء إنشائها، وارتفاع تكلفتها، وارتفاع سعر استضافتها مقارنة بالوثائق الثابتة.<sup>(4)</sup>

و. عادة ما يطلق عليها العنكبوتية العميقة (أو الخفية) deep web.

ومن أمثلة هذا النوع الوثائق المخزنة بمراصد البيانات والاستبانات واستطلاعات الرأي الإلكترونية.

### ثالثاً: الوثائق النشطة Active Documents:

تقوم بتنفيذ شفرة برمجية code تصدر إليها على الحاسب الآلي الخاص بالمستفيد. مثل: تحديث العرض update display، والاتصال عن بعد remote connection، والنماذج الذكية smart forms.<sup>(5)</sup>

ولا يقتصر الأمر على تعدد مجالات وفئات محتوى العنكبوتية العالمية، أو تسجيله على أكثر من نوع من الوثائق، ولكن تتعدد أيضاً اللغات المكتوب بها هذا المحتوى؛ بحيث تأتي اللغة الإنجليزية على رأس اللغات المستخدمة في كتابة محتوى العنكبوتية العالمية، تليها اللغات الأخرى، كالصينية والإسبانية واليابانية والفرنسية والألمانية، ثم تأتي اللغة العربية في المركز السابع.<sup>(6)</sup>

### 5. السمات العامة للنصوص العربية على العنكبوتية العالمية

هناك سمات وخصائص عامة تميز اللغة العربية قد لا تتوافر في لغة أخرى، منها:

أ. أن النظام الإملائي للغة العربية يقوم على ثمانية وعشرين حرفاً، بالإضافة إلى الأشكال الأخرى التي تتخذها بعض الحروف، مثل: الألف المقصورة (ى)، والتاء المربوطة (ة)، والهاء (ه)، والهمزة (ء).<sup>(7)</sup> فضلاً عن أن بعض الحروف العربية تحتوي على نقط؛ فهناك عشرة حروف تحتوي على نقطة واحدة: ب، ج، خ، ذ، ز، ض، ظ، غ، ف، ن، وثلاثة حروف تحتوي على نقطتين: ت، ق، ي، وحرفان يحتويان على ثلاث نقط (ث، ش).<sup>(8)</sup>

ب. اختلاف شكل الحرف حسب موقعه من المقطع. والمقطع هو: مجموعة من الحروف المتصلة، ويمكن للكلمة الواحدة أن تتكون من أكثر من مقطع، والحرف إما أن يكون: مستقل بذاته، أو في بداية المقطع، أو في وسط المقطع، أو في آخر المقطع. كما تتشكل

الحروف عند كتابتها؛ فمنها ما يأخذ شكل، ومنها ما يأخذ شكلين، ومنها ما يأخذ أربعة أشكال: (9)

- شكل واحد: الهمزة (ء) و هي الحرف الوحيد الذي يأخذ شكل واحد.
- شكلان: وهي الحروف التي لا توصل بما بعدها، ومنها حرف الواو الذي قد يأخذ الشكل: و، وقد يأخذ الشكل: —و.
- أربعة أشكال: وهي الحروف التي توصل بما قبلها و ما بعدها، ومنها حرف العين (ع) والذي قد يأخذ الأشكال، مثل: ع، عد، عـ، عـع.
- ج. يأتي حرف الهمزة أحياناً كحرف مستقل، وأحياناً أخرى يأتي غير مستقل: فإذا كانت على السطر فهي حرف مستقل، أما إذا كانت على حرف مد، مثل: أ إ آ و ئ ، فهي غير مستقلة.
- د. تقرأ و تكتب نصوص اللغة العربية من اليمين إلى اليسار ما عدا الأرقام من الجهتين.<sup>(10)</sup> فمثلاً 25 تقرأ من اليمين إلى اليسار و لكن 125 تقرأ من اليسار ثم اليمين ثم الوسط و هكذا...
- هـ. هناك حروف تكتب ولا تنطق، وحروف تنطق ولا تكتب. من الحروف التي تكتب ولا تنطق: همزة الوصل مثلاً في عبد الله، و اللام الشمسية، وألف واو الجماعة في قالوا مثلاً. ومن الحروف التي تنطق و لا تكتب الألف في لكن فهي تنطق لاكن و هذا تنطق هاذا و الرحمن تنطق الرحمان وهكذا...
- و. الكلمات العربية مركبة يدخل فيها: التذكير والتأنيث -الزمن -العدد -المخاطب - المتكلم ... الخ.
- ز. قد تكتب العربية بأنماط مختلفة من الحروف المستخدمة في الخطوط العربية: النسخ، الرقعة، الثلث، الكوفي، الديواني... فمثلاً حرف الهاء يكتب بأكثر من شكل في العديد من الخطوط: ه ه ه ه ه ، وغيرها من الحروف.<sup>(11)</sup>
- ح. الامتداد الأفقي لبعض الحروف العربية أعرض من البعض الآخر، فمثلاً حرف الصاد (ص) ليس بعرض حرف الألف (ا).
- ط. وجود تشابه كبير بين أشكال حروف اللغة العربية، فمثلاً التاء (ت) تشبه الثاء (ث)، والصاد (ص) تشبه الضاد (ض)، والباء (ب) تشبه الياء (ي).
- ي. وجود علامات الضبط بالشكل ( َ ِ ُ ً ٌ ٍ ً )<sup>(12)</sup> و هي تنقسم إلى:

- الحركات: الفتحة، والضمة، والكسرة، والسكون، والفتحتان، والضمتان، والكسرتان.

- الشدة. وتتميز الشدة بأنها تأتي على الحرف مع إحدى الحركات ( َ ِ ُ ً ٌ ٍ ً ٍ ً ٍ ) .

ك. الضبط بالشكل لا يظهر أو يرد في غالب النصوص (ما عدا القرآن الكريم). وال ضبط بالشكل يؤثر على فهم الكلمة، فمثلاً: كلمة جمل يمكن أن تكون جَمَل أو جُمَل لذلك نحن نستعين بالسياق لمعرفة الكلمة.

ل. اتصال الكلمة بحروف السوابق واللواحق:

- السوابق: فمثلاً كلمة "الإسلام" يمكن أن تكون "بالإسلام" أو "للإسلام"، فتكون الباء واللام وأل التعريفية زوائد.

- اللواحق: مثل نون التوكيد، في: لتأكلن.<sup>(13)</sup>

م. تتكون أصغر وحدة للجملة في أبسط صورها من اسمين، مثل: "الله أحد"، وربنا الله، وقد تتكون من فعل واسم، مثل: انتبه، وفاز المُجد، أو من اسمين وحرف، مثل: في البكور البركة. ومن المستحيل أن تتكون الجملة من حرف وفعل فقط، أو من حرف واسم فقط، إلا إذا كان هناك مُقدر، كما في: "لم يلد"، ويا خفي الألفاظ.<sup>(14)</sup>

ن. يتفاوت إملاء الكلمات ورسمها بحركات التشكيل في النصوص العربية؛ ومرجع ذلك هو الأخطاء البشرية، واختلاف الممارسات أو القواعد المتبعة لإملاء الكلمات المعربة والأسماء الأجنبية، مثل محاولة إيجاد مقابلات عربية لكلمات: تليفزيون، وتليفون، وتليجراف... الخ. إلى أن تم اعتماد تعريب هذه الكلمات في النهاية بسبب تأخر الوصول لترجمة مرضية أو بسبب الفشل في التوصل إلى اتفاق في هذا الشأن.<sup>(15)</sup>

س. قد تأتي المفاهيم على صيغ أفعال في الكثير من وثائق العنكبوتية العالمية، مثل: كيف تصلي؟ نريد بها الصلاة، كيف نتعلم؟ نريد بها التعليم. فالكلمات المفتاحية هنا هي المفاهيم المشتقة من الأفعال.<sup>(16)</sup>

ع. تستخدم الأرقام العربية الغبارية ( 1 2 3 4 ... ) في المشرق العربي، بينما تستخدم الأرقام العربية ( 1 2 3 4 ... ) في المغرب العربي وأوروبا.<sup>(17)</sup> على الرغم من الجدل الناتج حول أيهما أفضل.

## 6. ضوابط تشكيل المحتوى العربي المكتوب

هناك مجموعة من الضوابط المقترحة في تشكيل المحتوى العربي النصي، والتي قد تتفق فيها اللغة العربية مع لغة أو أخرى. ومن شأن هذه الضوابط أن تيسر استرجاع المحتوى العربي على العنكبوتية العالمية إذا ما تم الالتزام بها. وهذه الضوابط هي: (18)

أ. تشكل جميع الحروف بما في ذلك واو العطف و (لا) وبقيّة الحروف، ويستثنى من ذلك ما سيرد ذكره.

ب. الألف (ا): يشكل فقط في بداية الكلام وبعد علامات الترقيم التالية: . ، ؟ ! ؛

ج. اللام في أداة التعريف "الـ" التعريفية: لا تشكل إذا كانت شمسية وتوضع عليها سكون إذا كانت قمرية.

د. رسم تنوين النصب: يسبق الألف دائماً ولا يأتي بعده أو عليه، مثل: علماً. (19)

ه. حروف المد (الألف والواو والياء الساكنة والمسبوقة بحركة من جنسها): لا تشكل، مثل: علوم.

و. المسافة بين الكلمات:

- عدم الفصل بين واو العطف والكلمة التي تليها، مثل: ومحمد. إلا أن ذلك يمثل مشكلة لنظم الاسترجاع بشكل عام؛ إذ أن واو العطف من الحروف التي ترد كثيراً جداً في اللغة العربية، سواء المنطوقة أو المكتوبة، وعلى الرغم من سهولة تمييزها دلاليًا إلا أنه من الصعب الفصل بينها وبين الكلمة التي تليها إملائيًا، حيث يتم اعتبارها جزءًا لا يتجزأ من الكلمة. ومما يزيد من الأمر صعوبة بدء الكثير من الكلمات العربية بواو أصلية، مثل: وثائق - وسائل. فاحتساب واو العطف جزءًا من الكلمة إملائيًا عند تكشيف المحتوى العربي على العنكبوتية العالمية سيقفل من كفاءة استرجاع محركات البحث.

- وضع مسافة واحدة بين جميع الكلمات بما في ذلك: أو، لا، عبد.

ز. عدم استخدام الكشيدة، كما في "محمــــــــــــد" بل تكتب: محمد.



## 7. الخصائص اللغوية للمحتوى العربي على العنكبوتية العالمية

أ. الاعتدال وحدة الخاصية الصرفية:

بنيت اللغة العربية على مبدأ الاعتدال؛ فأكثر كلماتها وضعت على ثلاثة أحرف، وقليل منها أصله رباعي أو خماسي، لئلا يطول النطق ويعسر. ونلاحظ عدم الإكثار من الألفاظ الثنائية؛ خشية تتابع عدة كلمات في العبارة الواحدة فيضعف متن الكلام ويحدث فيه ما يشبه التقطع لتوالي الألفاظ المكونة من حرفين.<sup>(20)</sup> كما تتسم اللغة العربية بوحدة خاصيتها الاشتقاقية؛ إذ أنها تتميز بالاطراد الصرفي شبه المنتظم، إضافةً إلى ظاهرة التعدد الصرفي، كتعدد صيغ الجمع، مثل: طلاب، طلبة.

ب. المرونة النحوية:

فاللغة العربية تتمتع بشيء من الحرية النسبية في ترتيب الكلمات داخل الجمل، ويتجلى ذلك في تراكيب الجمل بفعل آليات التحويل النحوي المختلفة: كالنقيد والتأخير (إنما يخشى الله من عباده العلماء)، والحذف والإبدال، كاستخدام اسم الفاعل بدلاً من الفعل (ضارباً أخاه) والمصدر بدلاً من الظرف (استيقظت طلوع الشمس).<sup>(21)</sup>

ج. الحساسية السياقية:

حيث تتسم اللغة العربية بنوع من تأخي العناصر اللغوية مع ما يحيطها أو يرد معها من عناصر. ويمكن اقتفاء هذه الحساسية السياقية على مستويات متعددة؛ فعلى مستوى الإملاء يتوقف شكل الحرف العربي على الحرفين السابق واللاحق له، مثل: شكل حرف العين في شعبان، عناصر، مزرعة، مربع. وعلى مستوى النحو تبدو مظاهر هذه الحساسية في علاقات المطابقة، كتطابق الصفة مع الموصوف، وتطابق الفعل مع الفاعل، والمبتدأ مع خبره المشتق، وإعراب الفعل مع أداة النفي السابقة له.<sup>(22)</sup>

د. شدة التماسك بين عناصر المنظومة اللغوية:

ويتجلى ذلك في:

- التكامل بين نظام الصرف ونظام النحو والمعجم؛ إذ ينعكس الاشتقاق بشكل واضح على تنظيم المعجم وأسلوب استخدامه. فضلاً عن أن نحو العربية وصرافها وجهان لعملة واحدة يصعب الفصل بينها.<sup>(23)</sup>

- شدة الصلة بين مباني الصيغ الصرفية ومعانيها، كارتباط صيغة "انفعل" بمعنى تطاوع وصيغة "تفاعل" بمعنى تبادل.

هـ. العبارات الطلبية:

بعض الجمل لا تعني ما تقوله حرفياً، ويتحدد معناها على معنى المقام، كأن نقول: هل من الممكن أن تعطيني القلم، فهذا ليس سؤالاً يحتاج لإجابة، ولكنه طلب. وكل أشباه هذه العبارات التي تعني غير مدلولها الحرفي تمثل تحدياً أمام نظم استرجاع المعلومات، التي يصعب أن تفهم المواقف الاجتماعية أو تفسرها أو تتصرف وفقاً لها. (24)

و. توافق اللفظ والمعنى:

تضع اللغة العربية أغلب الألفاظ والمسميات على اختيار صفة من صفات الشيء الذي يُراد تسميته، أو بعض أجزائه أو نواحيه، أو تحديد وظيفته وعمله، واشتقاق لفظ يدل عليه. فمثلاً تسمية الدراجة في العربية تشير إلى وظيفتها وعملها وحركتها. أما في الفرنسية فإن: bicyclette، تشير إلى أجزائها وتركيبها وحالتها الساكنة. ومثل ذلك السيارة التي تشير تسميتها إلى عملها بينما في الفرنسية كلمة automobile تعني المتحرك بنفسه. (25)

ز. التعريب والتوليد وكثرة استخدام المصطلحات الأجنبية:

يكثر استخدام المصطلحات الأجنبية المكتوبة بحروف عربية على العنكبوتية العالمية، وتتراوح هذه المصطلحات بين المصطلحات العلمية والأكاديمية التي تنعت موضوعات واسعة أو مفاهيم ضيقة، إضافةً إلى المصطلحات العامة المستخدمة في الإشارة إلى الأجهزة والآلات وغيرها. كما تغطي المصطلحات الأجنبية كافة فروع المعرفة، مع استخدامها بشكل أكبر في العلوم البحتة والتطبيقية.

وقد ساهم نمو وزيادة تفرع المعرفة في دخول المزيد من المفاهيم الجديدة إلى اللغة العربية، سواء استطاعت اللغة استيعاب المفهوم الحديث بألفاظ عربية أم لا. ويتجلى ذلك بوضوح في الدوريات العربية المتخصصة والمترجمات وغيرها، والتي يكثر بها المصطلحات الأجنبية دون أن يكون لها مقابلات عربية ثابتة، مثل: الميتاديتا، والميتادتا،

و الميتاداتا، ناهينا عن إشكاليات ترجماتها؛ هل هي ما وراء البيانات، أم البيانات الخلفية، أم بيانات عن البيانات...؟

ونجد لغة مثل اللغة الإنجليزية أدق وأشمل من اللغة العربية من حيث الدلالة وضبط المصطلحات، بما يسهم في دعم ووفرة إنتاجها الفكري. فلو أخذنا المثال السابق الذي ضربناه على مصطلح الـ: Metadata، لوجدنا ندرة كبيرة في مؤلفاته باللغة العربية، وتتركز معظمها على التطبيقات العامة لهذا الموضوع، إضافة إلى قلة مصطلحاتها أو مفاهيمها المركبة في اللغة العربية. بينما نجد آلاف الوثائق الإنجليزية على العنكبوتية العالمية حول الموضوع، ويقابلها غزارة في مصطلحاته أو مفاهيمه المركبة، والتي تغطي أيضاً كافة جوانب الموضوع وتفرعاته وتطبيقاته؛ سواء في المواقع الحكومية على العنكبوتية العالمية، أو الوثائق الأرشيفية، أو المتاحف، أو المواد السمعية والبصرية...

كما أننا نجد كثرة في استخدام أسماء لغات وبرامج الحاسب والعناصر الكيميائية والمصطلحات الطبية ومسميات الأجهزة والآلات، بل إن بعض المصطلحات الأجنبية قد تستخدم أكثر في البحث عن مقابلاتها العربية إن وجدت؛ ومرجع ذلك أن المصطلحات الأجنبية قد تكون أدق في استرجاع محتوى العنكبوتية العالمية وأكثر دلالة على المفهوم المبحوث عنه من وجهة نظر المستفيد.

ومن المشكلات التي تعاني منها اللغة العربية حال استخدامها لمصطلحات أجنبية، اختلاف الإملاء والهجاء (إلكترونيات، ألكترونيات، كمبيوتر، كومبيوتر)، وتركيب المفاهيم واشتقاقاتها ودمجها بكلمات عربية أصيلة (الديناميكا الحيوية)، إضافة إلى اختلاف درجة انتشار المصطلحات، وكذلك رسمها، من منطقة جغرافية لأخرى، وفي استخدام حروف العلة أو تركها واختلاف النقحرة مع بعض الحروف، مثل: ج، غ، وس، ث، ز، وت، ط. (البيسك، البيزك، والديموجرافيا، الديموغرافيا) وغيرها الكثير. (26)

وفي هذا السياق تجدر الإشارة إلى أن التعريب هو إحدى الوسائل التي يمكن بها السيطرة على المصطلحات الأجنبية؛ إذ ما اتبعت قواعده، بحيث يتم نقل الكلمة الأجنبية كما هي مع شيء من التعديل في صورتها، وبحيث تتماشى مع البناء العام والقواعد الصوتية والصرفية للغة العربية،

مثل لفظة: إبريق، وسندس، وزنجبيل، وتلفاز، وغيرها من الألفاظ غير عربية الأصل. (27)

وقد يتم التعريب بالشكل العشوائي الذي يؤدي إلى ابتداع المجتمع أو صكه لمصطلح جديد، ككلمة التلفزيون مثلاً، أو يتم بطريقة ممنهجة عن طريق مجامع اللغة العربية مثلاً، ويوجد في الوطن العربي عدة مجامع للغة العربية تختلف في موقفها من تعريب المصطلحات مما يؤدي إلى حدوث لغط كبير في أوساط مستخدمي هذه المصطلحات. فهي قد تكون معربة بشكل حرفي لدرجة أنها تفقد معناها التقني أو قد تكون مبنية على فهم خاطئ للمصطلح الأجنبي.

وكما نرى أن ألفاظ اللغة العربية ترتبط في نسق خاص في حروفها وأصواتها، ومادتها وتركيبها، وهيئتها وبنائها، وحينما تدخل لفظة أعجمية على اللغة العربية فإنه من الأفضل أن تسير على أوزان العربية وهيئتها وصيغها لكي تصبح عضواً كامل العضوية في الأسرة اللغوية. فقديماً كانت الألفاظ قليلة محدودة، إلا أنها كثرت في الآونة الأخيرة، مثل أسماء الألبسة والأطعمة والنباتات والحيوان وشتون المعيشة أو الإدارة... وغيرها، بسبب الانفتاح الكبير على المجتمعات، وتوافر أدوات الوصول الحر للمعلومات.

ويقوم التعريب على إحدى طريقتين:

- تغيير حروف اللفظ الدخيل، وذلك بنقص بعض الحروف أو زيادتها مثل:

**B physics** فيزياء/فيزيكا ، و **B geology** جيولوجيا

أو إبدال حرف عربي بالحرف الأعجمي:

براديس **B** فردوس

- تغيير الوزن والبناء حتى يوافق أوزان العربية ويناسب أبنيتها، وذلك بنقص بعض الحروف أو زيادتها أيضاً، وتغيير مدوده وحركاته، مع مراعاة سنان العربية الصوتية كمنع الابتداء بساكن، ومنع الوقوف على متحرك، ومنع توالي ساكنين...

وأكثر ما بقي على وزنه وأصله من الألفاظ هو من الأعلام: سجستان،

رامهرمز..

أما دليلهم إلى معرفة الدخيل فهو إحدى ثلاث طرق:

- فقدان الصلة بينه وبين إحدى مواد الألفاظ العربية:

بستان: ليس في العربية مادة بست.

- أن يجتمع فيه من الحروف ما لا يجتمع في الكلمة العربية:

ج ق جوسق - ج ص جَصَّ - ج ط طازج...

- أن تكون على وزن ليس في العربية:

بُرَيْسَمَ إفعيل - آجر فاعل... (28)

## 8. النص

النص ليس سلسلة من الكلمات والجمل وال فقرات، بل هو بنية معقدة متعددة المستويات، وشبكة كثيفة من علاقات الترابط اللغوي والدلالي والتماسك المنطقي. والجملة ليست حصيلة معاني ألفاظها، ومعنى النص ليس مجرد تجميع معاني جملته المتتالية. فالنص ساحة رمزية يتداخل فيها السياق على مختلف مستوياته من الإحياء النفسية والاجتماعية والتاريخية والفكرية. كما أن النص لا تحده بدايته ولا نهايته، ولا يمكن رسم حدود فاصلة بين خارجه وسياقه، ومن المستحيل رده إلى جذر أو معنى؛ فكل نص يرث مباني ومعاني مما سبقه من نصوص تسري بداخله وتمنحه قراءة ومعاني جديدة. (29)

## 9. المفهوم

المفهوم هو المعنى الدال الملازم للكلمات التي يتألف منها النص. والكلمة في المعنى اللغوي هي اللفظ المفرد، أما في اصطلاحها المعلوماتي فهي أصغر وحدة في اللغة تستطيع نقل معنى محدد في ذاتها، وهي قادرة على البقاء كوحدة مستقلة في جملة. أما المصطلح فهو كلمة أو عبارة تستعمل للدلالة على مفهوم.

وتحتاج المفاهيم في عملية الاتصال إلى لغة واضحة أو كلمات موحدة بين المرسل والمتلقي حتى يمكن تخزينها ومعالجتها واسترجاعها. وتعتمد جل محركات البحث على اللغة الطبيعية في الوصول إلى المعلومات، وذلك بتكسييفها للمفردات التي تشكل بمجملها نصاً يحتوي على كلمات دالة وأخرى غير دالة.

ولذلك فإن المفردات غير الدالة تعيق الفهم والوصول السريع إلى المعلومات أثناء البحث على العنكبوتية العالمية. أما تقييم أهمية المفاهيم فيتم بالنظر إليها من زاويتين: الزاوية اللغوية، والزاوية المعلوماتية أو احتياجات المستفيدين حسب إجراءات وسياسة التكشيف المتبعة لدى محرك البحث.

ولذا لابد من التمييز بين النص الذي يتم تخزينه ومعالجته داخل نظام مغلق وموجه للتعامل مع المفردات داخل سياقاتها الضيقة أثناء الاسترجاع، وبين المفاهيم التي لها دلالات ثابتة نسبياً ومستقلة ضمن الإطار العام لنظام المعلومات الموجه لفئات عريضة من المستفيدين على العنكبوتية العالمية.<sup>(30)</sup>

## 10. النظام العددي في اللغة العربية

يتميز النظام العددي العشري بعدة مميزات، يمكن إيجازها فيما يلي:<sup>(31)</sup>

أ. استخدام تسعة أرقام والصفير:

يمكن كتابة أي عدد، مهما كانت قيمته، باستخدام تسعة أرقام والصفير، سواء أكان عدداً صحيحاً أم كسراً. ولسنا في حاجة لمعرفة أكثر من هذه الرموز العشرة لنبني منها الأعداد مهما بلغت قيمتها. إلا أن العربية تعاني من مشكلة الصفير الذي يُرسم بطريقة الأرقام الهندية المستعملة حالياً بهيئة نقطة، وكثيراً ما يؤدي تناهياها في الصغر إلى الوقوع في الخطأ. مما يُلتبس بينه وبين النقطة علامة الوقف أو النقط.

ب. الترتيب والقيمة المكانية للرقم:

يتميز النظام العددي الحالي بترتيب ثابت لا يتغير: فمثلاً تأتي 4 قبل 5، و5 قبل 6، و8 بعد 7، وهكذا في أرقام الأحاد، والعشرات، والمئات.. الخ، وهذا الترتيب الموضوعي للرقم في غاية الأهمية؛ لأنه يبين قيمة العدد عند مقارنته بغيره من الأعداد، ويبين تدرج الأعداد. وتختلف قيمة الرقم تبعاً لوضعه المكاني في العدد، فمثلاً العدد 4444 مكوّن من رمز واحد متكرر، وهو 4، لكن قيمة الـ 4 تختلف حسب مكانها، فالأولى 4 آحاد، والثانية 4 عشرات، والثالثة 4 مئات، والرابعة 4 آلاف. وما ينطبق على هذا الرقم ينطبق على غيره من الأرقام في مختلف الأعداد. ولا ريب أن الصفير هو الذي سهل استخدام الخاصية المكانية للرقم، وسهل كتابة الأعداد، وإجراء العمليات

المختلفة؛ فالعلاقة مثلاً -بين 5 و 6 تعطي نفس المعنى بين 50 و 60، وبين 500 و 600 وهكذا.

ج. الأساس 10:

أساس النظام العددي الحالي هو العشرة، أي أننا بعد الأرقام التسعة الأولى نكوّن وحدة من عشرة، ونجمع بعد ذلك في وحدات من عشرات.<sup>(32)</sup>

## 11. الألف واللام في أوائل الكلمات الإملائية العربية على العنكبوتية العالمية

هناك نوعان من الألف واللام في أوائل الكلمات الإملائية في اللغة العربية:

أ. الألف واللام التعريفيتان:

- برزت الألف واللام التعريفيتان بصعوبات جمة يمكن إجمالها في النقاط التالية:
- الألف واللام التعريفيتان زائدتان من حيث كونهما لا يحملان أية دلالات مهمة للاسترجاع الآلي.
- الألف واللام التعريفيتان مكونة من حرفين هما الألف واللام اللذان يلتصقان دائماً بأوائل الكلمات العربية؛ بما ييسر تمييزهما آلياً بسهولة، مع الأخذ في الاعتبار الألف واللام الأصليتين.
- توجد الألف واللام التعريفيتان بكثرة في الكلمات العربية الدالة، ومن النادر أن نجد محتوى عربي خالياً منهما.
- احتساب الألف واللام التعريفيتان في ترتيب الكلمات المفتاحية يؤدي إلى تشتتها كمدخل تكشفية؛ فقد توزع الكلمة تحت الألف واللام، ثم توزع مرة أخرى تحت الحرف الثالث بدون الألف واللام التعريفيتان.
- يؤدي احتساب الألف واللام التعريفيتين في الترتيب إلى تكديس أغلب الكلمات في موضع واحد، كما يؤدي إلى ضرورة إدخالها كل مرة يتم فيها المستفيد بعملية البحث، وقد يجعله يفترض أن مصطلح البحث قد يرد معرفاً ب الألف واللام أو قد يكون خالياً منهما، مما يزيد العبء على المستفيد ويقلل من كفاءة الاسترجاع.

ويمكن معالجة الألف واللام التعريفيتين آلياً كالتالي:

- إهمال الألف واللام التعريفيتين من منظور نظم الاسترجاع.

- ترتيب الكلمات المفتاحية في موضع واحد سواء أكانت بالألف واللام أو بدونهما.
- استرجاع مصطلحات البحث سواء أكانت بالألف واللام أو بدونهما.
- عدم التأثير على كتابة الكلمات ذات الألف واللام التعريفيتين وبقاؤها متصلة بهما.
- التمييز بين الألف واللام التعريفيتين والألف واللام الأصليتين.
- المعالجة الفردية للكلمات القليلة الملتبسة.

ب. الألف واللام الأصليتان:

تمثل الألف واللام الأصليتان جزء لا يتجزأ من الكلمة؛ فلا يمكن أن نبتزهما بحيث يستقيم معهما السياق، مثل: الله، آله، إلكتروني... أما الألف واللام التعريفيتان فهما زائدتان يمكن حذفهما دونما أي تغيير في دلالة الكلمة. وحينما تدخل الألف واللام التعريفيتان على الألف واللام الأصليتين فإنهما يغيرا من شكل الحرف، مثل: الألف، الألف. مع الأخذ في الاعتبار المشترك اللفظي لبعض الكلمات، مثل: الغاز وألغاز. ولذا يصبح من الضرورة وضع قائمة استثناء تضم كافة الكلمات المبدوءة بالألف واللام الأصليتين عند التعامل الآلي من المحتوى العربي على العنكبوتية العالمية.<sup>(33)</sup>

## 12. اللهجات والأساليب اللغوية العربية على العنكبوتية العالمية

تفرعت من اللغة العربية الكثير من اللهجات والتي يكاد يصبح بعضها غير مفهوم بين أفراد الأمة العربية؛ وحدث ذلك بسبب التباين الاجتماعي من مجتمع عربي لآخر. ففي اللهجة المصرية، والتي تتسم بغزارة المفردات العامية، نجد أنه على الرغم من الإلغاء الرسمي للألقاب فإنه مازالت تستخدم بعض التعبيرات للخطاب والمناداة أو سابقة للاسم الرسمي أو المنصب الرسمي، مثل: يا سيد، يا أستاذ، يا دكتور، يا باشا، يا عم... وغيرها الكثير. مما يجعل الموضوعات تنشئت تبعاً لاختلاف اللهجات المستعملة من قطر لآخر. والشكل التالي يوضح أبرز اللهجات العربية المستخدمة.

كما أن الأساليب اللغوية العربية التي يستخدمها المستفيد من العنكبوتية العالمية تتغير بتغير الموضوع من جهة، والمشاركين في هذا الموضوع من جهة أخرى. كما تتغير بفعل عوامل أخرى قد تكون اجتماعية أو نفسية أو كلاهما معا. وتنقسم الأساليب العربية المستخدمة على العنكبوتية العالمية إلى خمسة أساليب، هي:



أ. الأسلوب الجامد: وهو الذي يستعمل فيه كلام رسمي جدًا لدرجة اعتبار المستفيد غير موجود لأنه لن يستطيع أن يؤثر فيما يقال أو يكتب، مثل: الخطب الرسمية، والأدعية، والصلوات، والإعلانات، وتلاوات الكتب المقدسة، والشعر، والمسرحيات. فالمستفيد هنا يتسم بالسلبية فيما يتلقاه.

ب. الأسلوب الرسمي: وهو أسلوب لا يشارك فيه المستفيد أيضًا، كالبرامج الإذاعية والمحاضرات.

ج. الأسلوب الاستشاري: وهو أسلوب فيه كثير من الأسلوب الرسمي، ولكنه يتطلب تفاعل المستفيد، ولا يخطط له المسئول عن المحتوى بعناية بالغة، كما يتجنب فيه استعمال العبارات المقننة أو الكليشوهات، بل يمكن أن يستعمل فيه بعض التراكيب والأشكال اللغوية المختصرة التي تستعمل في الحديث العادي.

د. الأسلوب العادي: وهو الأسلوب الذي يستعمله المستفيدون العاديون فيما بينهم، مثل جماعات النقاش discussion groups. وهذا الأسلوب يفترض خلفية من المفاهيم والمعلومات المشتركة بين المستفيدين، ويستعمل فيه كثيرًا من اللغة العامية، كما تستعمل فيه اللغة الخاصة بمهنة أو فئة معينة من المستفيدين إذا كان المشاركون في النقاش ينتمون إلى مهنة أو رابطة اجتماعية معينة. وهذا الأسلوب تكثر فيه المشاركات والمقاطع والتعليقات، كما تكثر فيه الجمل الناقصة وأشباه الجمل والإجابات بكلمات مفردة.

ه. أسلوب الألفة الشديد: وهو أقل الأساليب رسمية، ويتألف عادة من أشباه جمل ومفردات، وتستعمل فيه عادة اللغة العامية جدًا، بل كثيرًا ما تستخدم فيه عبارات خاصة بشريحة صغيرة جدًا. ويبدو أن وظيفة هذا الأسلوب تختلف عن وظائف الأساليب الأخرى وكان المقصود منه التعبير عن الأحاسيس أو العواطف أكثر من نقله للأفكار والمعلومات.

وبالرغم من هذه الأنواع فإنه كثيرًا ما يحدث الانتقال من أسلوب لغوي إلى أسلوب لغوي آخر في نفس المحتوى، لكي يؤدي الأسلوب الجديد وظيفة أخرى مغايرة للأسلوب المستخدم.<sup>(34)</sup>

### 13. النص الفائق Hypertext داخل محتوى العنكبوتية العالمية

مفهوم النص الفائق:

لم يعد المعنى الواحد للفظ متداولاً لجميع قرائه، والعكس بالعكس؛ فالكلمات كما ترد على العنكبوتية العالمية مجرد نقطة انطلاق لإبحار لا نهائي من البحث عن معنى أو موضوع. فقد هاجر المعنى بعيداً عن استاتيكية المعاجم إلى ديناميكية النصوص، ولم يعد المعنى الواحد متعددًا بذاته بل أيضًا متعددًا في قراءاته وأساليبه تلك القراءات. وأصبح النص لا يشير إلى جوهر بداخله بل يشير هو الآخر إلى نص أو نصوص أخرى في سلسلة لا نهائية من التناص *intertextuality*.

وقد تفجر النص في عدد لا نهائي من القراءات الممكنة وأصبحت مهمة إلحاق المعنى به مسئولية المستفيد لا كاتب المحتوى وحده. ولم يعد النص المكتوب على العنكبوتية العالمية مجرد أثر ملقى على سطور وثائقها، بل أصبح ذلك الكل المتداخل أو اللامتاهي الذي تتدرج تحته جميع احتمالات قراءة النص والتناص.<sup>(35)</sup>

إن تقنية النصوص الفائقة هو الأسلوب الذي يتيح للمستفيد تتبع مسارات العلاقات الداخلية بين كلمات النص وجمله وفقراته، ويخلصه من قيود خطية النص؛ حيث يمكنه من التفرع من أي موضع داخله إلى أي موضع لاحق أو سابق، بل يساعد أيضًا في عملية تكشيفه، بحيث يربط بين عدة مواضع في النص ربما تكون مترادفة أو مترابطة تحت كلمة أو عدة كلمات مفتاحية.

كما أن تقنية النصوص الفائقة لا تنظر إلى النص كسلسلة متلاحقة من الكلمات؛ بل كشبكة من علاقات التداخل، فقد أخذتنا من النصوص إلى التناص ومن الخطية إلى التشعب. الأمر الذي جعل البعض ينظر إليها على أنها نظامًا لإدارة الروابط *link management system* لا نظامًا لإدارة البيانات *data management system* كما هو الحال في مرصد البيانات التقليدية.<sup>(36)</sup>

خصائص يجب توافرها في النصوص العربية الفائقة:

هناك خصائص أو شروط إذا ما توافرت في النصوص العربية الفائقة، فإنها تؤدي بالقطع إلى تيسير استرجاعها؛ فالكثير من محركات البحث على العنكبوتية العالمية تنظر بعين الاعتبار إلى هذه النصوص والأسلوب الذي كتبت به. وهذه الخصائص هي:<sup>(37)</sup>

أ. كتابة النص العربي بصورته السياقية الطبيعية وكأنه نص غير فائق. مثال:

"... إن كتابة النص العربي الفائق وربطه بوثيقة أخرى طريقة مثلى لتحسين

[الترتيب الطبقي](#) للمفردات أو النصوص داخل محركات البحث..."

ب. أن يصف النص العربي الفائق الغرض منه، لا الآلية المكتوب بها. مثال:

غير مفضل: "... [اضغط هنا](#) للاطلاع على المزيد..."

مفضل: "... يمكنك الحصول على معلومات أكثر حول [النص العربي الفائق](#)

في..."

ج. أن يكون النص العربي الفائق واضحاً ودالاً بحيث يمكن فهمه من خارج

السياق. مثال:

غير مفضل: "... يمكنك متابعة آخر الأخبار عن طريق الدخول إلى [موقع](#) قسم

المكتبات والمعلومات..."

مفضل: "... يمكنك متابعة آخر الأخبار عن طريق الدخول إلى [موقع قسم](#)

[المكتبات والمعلومات](#)..."

د. أن يتوقع المستخدم ما سيسفر عنه تتبع النص العربي الفائق؛ هل سينتقل إلى

وثيقة أم سيبدأ في تشغيل ملف صوتي... الخ. مثال: "... يمكن [تحميل المدقق](#)

[اللغوي](#) مجاناً..."

إلى جانب نظم النصوص الفائقة هناك نظم أخرى لا غنى عنها للحرث المتعمق

في متن النصوص، مثل: النظم المبنية على المعرفة knowledge-based systems

التي تكفل تحويل النصوص إلى شبكات دلالية semantic nets أو مخططات مفاهيمية

conceptual graphs بحيث يسهل النفاذ إلى تفاصيلها الدقيقة بصورة منهجية منتظمة،

وهو ما أصبح يطلق عليه في الآونة الأخيرة: العنكبوتية الدلالية semantic Web.

ولكن المحتوى العربي على العنكبوتية العالمية يعاني من قلة ارتباطاته، بحيث

انتقلت خطية المحتوى التقليدي إلى تلك الشبكة، وهو ما يتناقض جوهرياً مع طبيعتها

المصممة أساساً للملاحة غير المحدودة، مما يؤدي -لا ريب- إلى انغلاق وانعزالية هذا

المحتوى الإلكتروني. وترجع قلة هذه الارتباطات إلى:

أ. غياب الفهم الواضح لخدمات المعلومات المقدمة من خلال العنكبوتية العالمية.

ب. نقص الدراسات البيئية.

ج. ضعف نظم تكشيف المحتوى العربي على العنكبوتية العالمية.<sup>(38)</sup>

## 14. الوضع الراهن للمحتوى العربي على العنكبوتية العالمية

أولاً: الجوانب الخاصة بمصادر المحتوى:

أ. فقر المحتوى: يعيب معظم المواقع العربية عدم ثراء محتواها وافتقار هذا المحتوى للتنوع، ولا يستثنى من ذلك سوى عدد محدود من المواقع، وهي مواقع المؤسسات التي تملك في الأصل محتوى مطبوع، مثل: دور الصحف والمواقع الدينية وهيئات الآثار، أو المؤسسات التي كانت سباقة في استخدام الحاسبات الآلية؛ كالبنوك وشركات الطيران وما شابه.

ب. طغيان النصوص: تغطي فئة النصوص على غيرها من فئات المحتوى السمعية والبصرية، وغالبًا ما تنقل هذه النصوص كما هي من أصولها المطبوعة دون تطويع لما تتطلبه تدابير الرقمنة. كما أن معظم النصوص ذات أصول عربية، حيث تندر النصوص المترجمة لضعف التركيز على المجالات العلمية -التكنولوجية ومعارف الفكر الحديث بصفة عامة والتي تعتمد أساساً على المصادر الأجنبية.

ج. غياب مفهوم إعادة استخدام مصادر المعلومات: إذ تعتمد صناعة المحتوى بصورة أساس على إعادة استخدام مصادر المعلومات، واتباع نظم حفظ أرشيفي فعالة، أو استغلال المعارف الناشئة عن التراكم المعلوماتي، وهو ما يتيح فرصاً متعددة ومبتكرة لم يكن في الإمكان توافرها فيما مضى لإعادة توظيف مصادر المعلومات في سياقات وأطر جديدة. ويمكن إرجاع هذه الظاهرة السلبية لغياب مفهوم إعادة الاستخدام إلى عدة أسباب رئيسية، هي:

- تبيد كثير من مواد الأرشيف الصحفي والتراثي.  
- تخلف نظم الحفظ الأرشيفي المستخدمة من حيث التنظيم والتصنيف وآليات البحث.

- قلة الإنتاج الوثائقي في الدراسات الإنسانية والاجتماعية والعلمية والتقنية.  
- ضعف النزعة التاريخية historization.

د. بزوغ المدونات ومواقع المشابكة المجتمعية بمعدلات متسارعة لتضيف بذلك شقاً معلوماتياً رئيساً في إطار استرجاع المحتوى.

هـ. غياب القيمة المضافة added value: وذلك من خلال:

- تعزيز المحتوى بروابط النصوص الفائقة والوسائط الفائقة.

- تعزيز المحتوى بالكلمات المفتاحية في ما وراء البيانات.
- إضافة المراجعات والعروض reviews والتحليلات وآراء المستخدمين.
- ترجمة بعض فقرات المحتوى إلى اللغات الأخرى.
- إثراء المحتوى بالرسومات التخطيطية وقوائم النقاط المحورية bullets.
- إعداد معاجم الألفاظ وقوائم المصطلحات.
- التجاوب مع العنكبوتية الدلالية.<sup>(39)</sup>

### ثانياً: الجوانب الخاصة بإدارة المواقع وتشغيلها:

أ. قصور إدارة المواقع: تترك كثير من المواقع العربية دون إدارة فعالة تواصل على تطويرها وتعكف على خدمة المستخدمين منها؛ إما لنقص الكوادر المؤهلة، أو لغياب الموارد المادية الكافية، أو لغياب الوعي الكافي والفهم الصحيح لتقنيات إدارة المواقع لدى المؤلفين أو الناشرين، مما يؤدي إلى التدهور التدريجي للمواقع. وما أكثر المواقع التي تورط خرائطها فئات من خدمات المعلومات دون أن تستوفي مطالبها وتعجز عن تقديمها.

ب. عدم انتظام دورة التحديث: يعاني كثير من المواقع العربية من عدم انتظام دورة تحديث محتواها وخدماتها، باستثناء تلك المواقع التي تفرض عليها طبيعة نشاطها ضرورة الالتزام بالتحديث بصفة دورية منتظمة، مثل مواقع الصحف ومواقع خدمات النقل والبنوك وما شابه. وكثير من تصميمات المواقع تغفل مطالب التحديث ولا تدفع بأدواته وطرائقه في بنية وتنظيم الموقع، ولا تقوم بحساب مؤشرات كمية دورية لقياس مدى تقادم المعلومات. ويؤدي عدم انتظام التحديث إلى جمود المحتوى إلى حد يفقده تميزه عن سابقه المطبوع، الأمر الذي يؤثر على جودة النتائج التي تقدمها محركات البحث.

ج. تخلف أساليب التصميم: يتسم تصميم أغلب المواقع بالبدائية، خاصة تلك المواقع التي تم إنشاؤها في مرحلة مبكرة ولم يتم تحديثها. وكثيراً ما يوكل تصميم المواقع إلى أشباه المحترفين الذين عادة ما يأتون بما يمكن أن نطلق عليه تصميمات سابقة التجهيز.<sup>(40)</sup> وتبرز شواهد تخلف التصميم في:

- ندرة استخدام نظم إدارة المحتوى (CMS) Content Management Systems، والتي توفر أدوات منهجية تكفل التحليل الدقيق لمسارات العمل داخل المؤسسات،

وترجمة ذلك إلى خرائط تدفق work flows تيسر الوظائف التي يقوم بها الموقع لدعم محتواه.

- افتقار الكثير من المواقع إلى خرائط site maps، ومنتبعات الأثر-bread crumb trails، والتي تيسر التصفح داخل المواقع وتسهل من التنقل بين وثائقه الرئيسية ووثائقه الفرعية.

- غياب الاطراد في تصميم الوثائق، وافتقار توزيع المحتوى داخل الوثائق إلى التناسق.

- ندرة الروابط الفاتقة داخل الوثائق وفيما بينها.<sup>(41)</sup>

## 15. لقاء اللغة العربية بتكنولوجيا الإنترنت

هناك مواجهة غير متكافئة بين اللغة العربية واللغة الإنجليزية على العنكبوتية العالمية بشكل عام؛ نظراً لطغيان الأساس الإنجليزي على توجهات العنكبوتية العالمية التكنولوجية، فقد حُددت هذه التوجهات أصلاً لتلبي احتياجات اللغة الإنجليزية بالدرجة الأولى، ويظهر ذلك في:

أ. تصميم معظم لغات البرمجة باللغة الإنجليزية.

ب. استخدام معايير وشفرات لتبادل البيانات مصممة في الأساس للتعامل مع الهجائية الإنجليزية المحدودة في عدد حروفها وأشكال هذه الحروف.

ج. تصميم معظم نظم استرجاع المعلومات على أساس أن اللغة الإنجليزية هي لغة الهدف.

د. القسم الأكبر من مكتبة البرامج الجاهزة باللغة الإنجليزية.

ه. معظم الكتب والمراجع والدوريات والأبحاث... باللغة الإنجليزية.

## 16. تعريب النظم الآلية للتعامل مع المحتوى العربي

لا يعدو الموقف الراهن لتعريب النظم الآلية للتعامل مع المحتوى العربي على العنكبوتية العالمية في مجمله كونه نتيجة منطقية لأزمة اللغة العربية الحادة، وواقع متأزم لمركزنا التقني والعلمي، ويبدو ذلك جلياً في الملامح الرئيسية التالية:

أ. محاولة استيعاب اللغة العربية في نطاق هذه النظم الآلية المصممة أساساً للغة الإنجليزية؛ فتحت ضغط الدوافع العملية والتجارية لمصممي البرمجيات، تم إتباع طرق

أشبه بالمتعسفة لإخضاع اللغة العربية للقيود التي فرضتها خصائص اللغة الإنجليزية. وقد تم ذلك إما بالتحايل على هذه القيود الفنية على حساب المستفيد وكفاءة النظام، أو بالترخص في بعض خصائص اللغة العربية؛ كالحذف من أشكال الحروف وإغفال علامات التشكيل، ورداءة التعامل مع الكتابة العربية إدخالاً وإخراجاً.

ب. سطحية التعريب؛ فقد تركزت معظم الجهود حول اللغة العربية على مستوى الحرف دون الصرف والنحو والمعجم والسياق، وأصبح مفهوم تعريب النظم الآلية للتعامل مع المحتوى العربي مرادفاً لإدخال النصوص العربية من لوحة المفاتيح وطباعة النصوص العربية أو إظهارها على الشاشات المرئية.

ج. الاعتماد على الجهود الأجنبية؛ حيث نلاحظ أن معظم جهود تعريب النظم الآلية للتعامل مع المحتوى العربي ظلت تفرغ إلينا من خارج الوطن العربي، وخاصة الولايات المتحدة الأمريكية.

د. ندرة الأبحاث التي تتناول اللغة العربية وخصائصها ومعالجتها آلياً.

## 17. الخلاصة

وجدنا كيف يتنوع المحتوى العربي على العنكبوتية العالمية، وكيف تتعدد مجالاته، ورأينا إلى أي مدى هذا المحتوى متردي ولا يرقى إلى المستوى الذي هو متاح في كثيرًا من المحتويات غير العربية الأخرى. كما وجدنا أن خصائص المحتوى العربي والظواهر التي تؤثر فيه ودلالات الكلمات... الخ ذات تأثيرات بالغة على التعامل الآلي معها. فإذا لم يؤخذ ذلك كله بعين الاعتبار سيظل هذا المحتوى مهمشًا ومتقادمًا ومنسيًا. ويمكن القول أن حال المحتوى العربي على العنكبوتية العالمية هو مشكلة خطيرة لا تقل أهمية أو خطورة وتعقيدًا عن المشكلات الأخرى التي تواجهها أمتنا العربية. ومما لا شك فيه أن لغتنا العربية غير مرشحة على الإطلاق للاندثار أو الاختفاء، إلا أن المحتوى العربي على العنكبوتية العالمية مهدد بهذا المصير. وقد آن لنا أن نهئى محتوانا العربي على العنكبوتية العالمية لمتطلبات تلك البيئة الإلكترونية، بحيث نبعث الحياة فيه تنظيرًا وبحثًا واسترجاعًا.

1. اللجنة الاقتصادية والاجتماعية لغربي آسيا. المحتوى الرقمي العربي؛ الفرص والأولويات والتوجهات. نيويورك: الأمم المتحدة، 2005. متاح في:  
[http://www.escwa.un.org/divisions/ictd/workshop/forum\\_a/docs/Digital%20Arabic%20Content%20Opportunities,%20priorities%20and%20strategies.pdf](http://www.escwa.un.org/divisions/ictd/workshop/forum_a/docs/Digital%20Arabic%20Content%20Opportunities,%20priorities%20and%20strategies.pdf). تاريخ الزيارة: 2007/7/25.
2. نبيل علي. منظومة صناعة المحتوى العربية؛ التحديات والفرص ومناهل الحلول. اجتماع خبراء حول تعزيز المحتوى الرقمي العربي، بيروت يونيو 2003. بيروت: اللجنة الاقتصادية والاجتماعية لغربي آسيا (إسكوا)، 2003. متاح في: <http://www.escwa.un.org/wsis/meetings/3-5june/docs/13.pdf>. تاريخ الزيارة: 2007/10/21.
3. تشاودھري، ج. ح. الإنترنت وبحوث استرجاع المعلومات؛ مراجعة علمية موجزة. ترجمة حشمت قاسم، في كتابه: الاتصال العلمي في البيئة الإلكترونية. القاهرة: دار غريب، 2005. ص ص 114 - 115.
4. Static vs dynamic Websites - what's the difference?. available at: [http://www.edinteractive.co.uk/pages/news\\_detail.php?recordID=4](http://www.edinteractive.co.uk/pages/news_detail.php?recordID=4). visited on: 11/9/2007.
5. Yerkey, Neil. Active server pages for dynamic database web access. Library Hi Tech, vol. 19, no. 2 (2001). p.p 133 -142. available at: <http://www.emeraldinsight.com/10.1108/07378830110394277>. visited on: 15/7 2007.
6. Top ten languages used in the Web. available at: <http://www.internetworldstats.com/stats7.htm>. visited on: 10/8/ 2008.
7. Al-Kharashi, Ibrahim A. A Web search engine for indexing, searching and publishing Arabic bibliographic databases. available at: <http://www.isoc.org/inet99/proceedings/posters/085/index.htm>. visited on: 23/8/2007.
8. Ahmed, Pervez; Al-Ohali, Yousef. Arabic character recognition; progress and challenges. Journal of King Saud University, vol. 12, (2000). available at: <http://www.ksu.edu.sa/printpress/researches/V12M152R682.pdf>. visited on: 7/4/2008.
9. Wegener, Sibylle. Arabic and the WWW. available at: [http://ssgdoc.bibliothek.uni-halle.de/vlib/html/docs/arabwww\\_en.html](http://ssgdoc.bibliothek.uni-halle.de/vlib/html/docs/arabwww_en.html). visited on: 15/5/2008.
10. Ibid.
11. نبيل علي. العرب وعصر المعلومات. الكويت: المجلس الوطني للثقافة والفنون والآداب، 1994. (عالم المعرفة؛ 184). ص 351.
12. محمد زكي محمد خضسر. الحروف العربية والحاسوب. مجلة مجمع اللغة العربية الأردني. ص 6. متاح في: <http://www.al-mishkat.com/khedher/Papers/paper4.htm>. تاريخ الزيارة: 2007/2/2.
13. Khoja, Shereen; Garside, Roger; Knowles, Gerry. A tagset for the morphosyntactic tagging of Arabic. Lancaster: Lancaster University, 2001. available at: <http://archimedes.fas.harvard.edu/mdh/arabic/CL2001.pdf>. visited on: 15/5/2007.
14. لقاء مع الأستاذ الدكتور شعبان صلاح الأستاذ بكلية دار العلوم -جامعة القاهرة. أغسطس 2008.
15. Kasem, Hishmat M. A. The performance of Arabic in mechanized information retrieval system; theoretical background and experimental design (drawn from the author's Ph.D. Thesis: Arabic in specialist information systems; a study in linguistic aspects of information transfer. London: University of London, 1978). Arab Journal for Librarianship and Information Science. vol. 3, issue 3. (July 1983). p. 23.
16. عبد الله شرف الغامدي؛ بدرية سليمان الفهود. أداة وب معتمدة على عملية التحليل الهرمي للحصول على معجم عربي موحد لتقنية المعلومات. الندوة الدولية الأولى عن الحاسب واللغة العربية، 2007. متاح في: <http://www.iscal.org.sa/joomla/images/proceedings/alghamdi2.pdf>. تاريخ الزيارة: 2007/9/25.
17. أحمد أبو الهيجاء. تقييس استخدام اللغة العربية في الحاسوب. اجتماع خبراء حول تعزيز المحتوى الرقمي العربي، بيروت يونيو 2003. بيروت: اللجنة الاقتصادية والاجتماعية لغربي آسيا (إسكوا)، 2003. متاح في: <http://www.escwa.un.org/wsis/meetings/3-5june/docs/13.pdf>. تاريخ الزيارة: 2007/11/1.
18. لقاء مع الأستاذ الدكتور إبراهيم الضوه الأستاذ بكلية دار العلوم -جامعة القاهرة. أغسطس 2008.
19. سمير شريف ستيتية. رؤية جديدة في تفسير التنوين في العربية. مجلة جامعة الملك سعود؛ الآداب، مج5، ع1 (1993). ص ص 119 - 144. متاح في: <http://www.ksu.edu.sa/printpress/researches/V5M83R2065.pdf>. تاريخ الزيارة: 2008/5/16.
20. محمد صالح الشنطي. المهارات اللغوية؛ مدخل إلى خصائص اللغة العربية وفنونها. حائل: دار الأندلس، 1996. ص 52.



21. نبيل علي. العرب وعصر المعلومات. المصدر السابق. ص ص 334- 335.
22. نفس المصدر. ص 336.
23. Kasem, Hishmat M. A. Subject and linguistic morphology of the Arabic specialist literature; the linguistic situation as displayed in published literature design (drawn from the author's Ph.D. Thesis: Arabic in specialist information systems; a study in linguistic aspects of information transfer. London: University of London, 1978). Arab Journal for Librarianship and Information Science. vol. 4, issue 2. (April 1984). p.4.
24. أبو الحجاج محمد بشير. تحديات تواجه معالجة الآلية للغة. لغة العصر، ص 8، ع 92 (أغسطس 2008). ص 17.
25. فرحان السليم. اللغة العربية ومكانتها بين اللغات. متاح في:  
<http://www.saaid.net/book/open.php?cat=90&book=2625> . تاريخ الزيارة: 2007/11/6.
26. نفس المصدر. ص ص 51- 54.
27. علي الطالقاني. في دائرة الاستهداف؛ اللغة العربية مخاوف من اندثارها. متاح في:  
<http://www.arabcin.net/modules.php?name=News&file=print&sid=1336> . تاريخ الزيارة: 2007/11/15.
28. فرحان السليم. اللغة العربية ومكانتها بين اللغات. متاح في:  
<http://www.saaid.net/book/open.php?cat=90&book=2625> . تاريخ الزيارة: 2007/11/6.
29. نبيل علي. الثقافة العربية وعصر المعلومات؛ رؤية لمستقبل الخطاب الثقافي العربي. الكويت؛ المجلس الوطني للثقافة والفنون والآداب، 2001. (عالم المعرفة؛ 265). ص ص 463.
30. علي السليمان الصوينع. استرجاع المعلومات في اللغة العربية. الرياض: مكتبة الملك فهد الوطنية، 1994. (السلسلة الثانية؛ 21). ص 34.
31. الأرقام. متاح في: <http://www.alargam.com/books/the%20numbers/18.htm> . تاريخ الزيارة: 2008/5/21.
32. نفس المصدر.
33. نفس المصدر. ص ص 122- 125.
34. نايف حرما. أضاء على الدراسات اللغوية المعاصرة. الكويت: المجلس الوطني للثقافة والفنون والآداب، 1978. (عالم المعرفة؛ 9). ص ص 184- 194.
35. Cicconi, Sergio. Hypertextuality. available at: <http://www.cisnet.com/cisnet/writing/essays/hypertextuality.htm>. visited on: 5/1/ 2008.
36. نبيل علي. الثقافة العربية وعصر المعلومات؛ رؤية لمستقبل الخطاب الثقافي العربي. المصدر السابق. ص 279.
37. How to write link texts that will increase your Google rankings. available at: <http://www.free-seo-news.com/newsletter329.htm#facts>. visited on: 12/6/ 2008.
38. نبيل علي. الثقافة العربية وعصر المعلومات؛ رؤية لمستقبل الخطاب الثقافي العربي. المصدر السابق. ص 100.
39. نبيل علي. منظومة صناعة المحتوى العربية؛ التحديات والفرص ومناهج الحلول. المصدر السابق. متاح في:  
<http://www.escwa.un.org/wsis/meetings/3-5june/docs/13.pdf> . تاريخ الزيارة: 2007/10/21.
40. Abdulla, Rasha A. Arabic language use and content on the Internet. Alexandria: Bibliotheca Alexandria, 2008. available at: <http://www.bibalex.org/a2k/attachments/references/rasha%20en.pdf> . visited on: 25/1/2008.
41. نبيل علي. مسح للمحتوى الرقمي وبرمجياته وتطبيقاته وتقييم احتياجاته. بيروت: اللجنة الاقتصادية والاجتماعية لغربي آسيا (إسكوا)، [د.ت.]. متاح في:  
<http://www.escwa.un.org/divisions/projects/dac/01DACNabi.pdf> . تاريخ الزيارة: 2007/11/26.