اللجنة الاقتصادية والاجتماعية لغربي آسيا

الارتقاء بجودة المحتوى الرقمي العربي

الأمم المتحدة
الاسكوا
ESCWA

اللجنة الاقتصادية والاجتماعية لغربي آسيا

الارتقاء بجودة المحتوى الرقمي العربي

الأمم المتحدة، بيروت
٢٠١٥

إخلاء مسؤولية:

صدرت هذه الوثيقة دون تحرير رسمي.

لم يتم التحقق من المراجع المذكورة في النص.

لا تنطوي التسميات المستخدمة في هذا التقرير، ولا طريقة عرض المادة التي يتضمنها، على الإعراب عن أي رأي كان من جانب الأمانة العامة للأمم المتحدة بشأن المركز القانوني لأي بلد من البلدان، أو أي إقليم، أو أية مدينة، أو أية منطقة، أو أي سلطة من سلطات أي منها، أو بشأن تعيين حدودها أو تخومها.

الآراء الواردة في هذه الورقة هي آراء المؤلف وليست، بالضرورة، آراء الأمانة العامة للأمم المتحدة.

لا يعني ذكر أسماء شركات أو منتجات تجارية أنّ الأمم المتحدة تدعمها.

كلمة شكر

قام بإعداد هذه الدراسة حول الارتقاء بجودة المحتوى الرقمي العربي السيد نور الدين شيخ عبيد، الأستاذ الجامعي والخبير في مجال تكنولوجيا المعلومات والاتصالات وتطوير مجتمع المعلومات.

وقد قامت السيدة نبال إدلبي، رئيسة قسم الابتكار، والسيدة هانيا صبيدين ديماسي، مساعدة أبحاث في قسم الابتكار في إدارة التكنولوجيا من أجل التنمية في الإسكوا بمراجعة هذه الدراسة وتدقيقها وإعدادها بالشكل النهائي.

وقد تم إصدار هذه الدراسة في إطار مشروع الإسكوا حول تعزيز صناعة المحتوى الرقمي العربي من خلال الحاضنات التكنولوجية الذي نفذته إدارة التكنولوجيا من أجل التنمية وتحت إدارة السيد حيدر فريحات.

المحتويات

مقدمة ١
أولاً - المحتوى الرقمي العربي وأدواته ٣
ألف- معطيات أولية ٣
باء- قضية المحتوى الرقمي ٥
جيم- المواقع الأولى في العالم العربي ٧
دال- مبادرات المحتوى العربي ٨
هاء- أدوات معالجة اللغات الطبيعية في خدمة المحتوى ١٠
ثانياً- معايير المواقع والمحتوى ١٣
ألف- المبادئ العامة المؤسسة لمحتوى المواقع ١٤
باء- معايير جودة المحتوى ١٦
جيم- معايير قياس نجاح الموقع ٢١
دال- عينة من مواقع عربية ومعايير الجودة ٢٢
هاء- دراسات في المحتوى الرقمي العربي ٢٤
ثالثاً- المحتوى ماضياً وحاضراً ومستقبلاً ٢٦
ألف- أجيال الوب ٢٦
باء- الوب الدلالي ٢٨
جيم- اللغة العربية والوب الدلالي ٢٩
رابعاً - التوصيات ٣٢
ألف- بالنسبة للدول العربية ٣٢
باء- الجامعات ومراكز البحوث والدراسات ومجامع اللغة العربية ٣٢
جيم- المنظمات الدولية والإقليمية ٣٣
دال- المجتمع المدني ٣٣
هاء- شركات تصميم وتنفيذ المواقع وصيانتها ٣٣

الملاحق

١- أبرز محاور البحث في معالجة النصوص والكلام والجهات المعنية ٣٧
٢- تقارير ودراسات الإسكوا في المحتوى الرقمي العربي ٣٩

مقدمة

أصبحت الإنترنت إحدى أهم وسائل تكنولوجيا المعلومات والاتصالات كونها تمثل الخازن الأساس والأكبر للمعلومات ووسيلة إيصالها إلى من يحتاجها. الأمر الذي جعل منها ذاكرة هائلة للمعلومات بخلاف الوسائط الأخرى التي تتميز بآنية معلوماتها كما هو الحال في الهواتف والتلفاز والإذاعة والمجلات والجرائد. وهي تتيح عبر بنيتها وأدواتها استرجاع المعلومات بسرعة فائقة لا يمكن للكتب والمراجع الورقية أن تضاهيها. وكذلك استرجاع كل ما هو رقمي من فيديو وصوت مما يتيح المشاهدة والاستماع إلى برامج في وقت غير وقت بثها.

قيل في بداية انتشار الإنترنت إن المحتوى هو الملك، وذلك للدلالة على أن الإنترنت هي محتواها، وأن فائدة الإنترنت كامنة في هذا المحتوى. وحينها، لم يكن المحتوى بالشيء الضخم كما هو اليوم، إلا أنه كان محط الاهتمام وجهود التطوير والإثراء المستمرة. وهذا المحتوى هو ما جعل مألوفاً رؤية الناس عموماً منكبين على أجهزة الاتصال كالهواتف الذكية، أو الحواسيب المحمولة واللوحية، يتداولون محتوى يرسلونه أو يستقبلونه أو يتعاملون معه.

عرَّفت الإسكوا المحتوى الرقمي العربي بأنه أي محتوى باللغة العربية بالشكل الرقمي سواء كان نصاً أو صورة أو فيديو، وهو يشمل المحتوى العلمي والاقتصادي والترفيهي والأدبي والإداري على مواقع الإنترنت، والخدمات الإلكترونية المختلفة كالحكومة الإلكترونية والتجارة الإلكترونية والتعليم الإلكتروني والصحة الإلكترونية، كما يشمل البرمجيات وقواعد البيانات ومنتجات المصدر المفتوح الداعمة والأدوات وبرامج معالجة اللغة العربية والمحركات البحثية ومحركات الترجمة. أما صناعة المحتوى الرقمي العربي، فتعرفها الإسكوا بأنها المخرجات الناتجة من أنشطة الشركات والمؤسسات الصغيرة والمتوسطة والناشئة العاملة في تطوير وإنتاج وتوزيع المحتوى الرقمي العربي؛ إضافة إلي الشركات التي تستضيف المحتوى الرقمي العربي أو معنية بتطوير المحتوى والبرمجيات التي تدعم وتعنى باللغة العربية[١].

لقد تنامى المحتوى على شبكة الإنترنت بشكل كبير منذ بداية هذا القرن، وتعبّر عن هذا التنامي معطيات تشير إلى أنه في كل دقيقة[٢] تُرسل أكثر من ٢٠٠ مليون رسالة بريد إلكتروني، ويتلقى غوغل أكثر من مليوني طلب بحث، وتُرسل أكثر من ١٠٠ ألف تغريدة على تويتر، وقرابة ٧٠٠ ألف مشاركة على الفيسبوك، وأكثر من ٢٧ ألف مشاركة جديدة على مدونات تمبلر (Tumblr)، ويجري تحميل أكثر من ٤٨ ساعة فيديو على موقع يوتيوب.

وبسبب طبيعة الإنترنت الشمولية والمتاحة للجميع ، تزايد المحتوى على نحو تراكمي مضطرد. والجديد في الأمر أن المحتوى أصبح، مع جيل الوب الثاني، يشارك في تأليفه المتصلون بالشبكة، ولم يعد المحتوى هو الملك بل أصبح كل متصل بالشبكة ملكاً بطريقة ما. أي أن تطور تكنولوجيات الإنترنت عزز المشاركة الفردية[٣] وكذلك مشاركة الجماهير وهي التوجهات التي من المتوقع أن تبقى بارزة في العقدين القادمين. أضف إلى ذلك التطور الدائم الذي يتسم به المحتوى، ليس فقط في

[١] الإسكوا. مذكرة سياساتية حول المحتوى الرقمي العربي. ٢٠١٣، E/ESCWA/ICTD/2013/Technical paper.5

[٢] http://www.domo.com/blog/2012/06/how-much-data-is-created-every-minute/?dkw=socf3

[٣] http://www.dni.gov/index.php/about/organization/global-trends-2030

مضمونه وتطبيقاته، وإنما في طريقة التواصل مع المستفيد منه وأدوات تقديمه، حيث التطور التكنولوجي المستمر. فمن المحتوى النصي إلى المحتوى التواصلي، ومن ثم إلى المحتوى الدلالي، فالمحتوى الانفعالي حيث سيكون للمستخدم إمكانية التخاطب مع وسيلة النفاذ إلى الإنترنت.

يوفر النفاذ إلى الإنترنت فرصاً كبيرة في إحراز تقدم في أعمال التفكير والإبداع عبر تخطي أعباء التذكر والحفظ إضافة إلى الاستزادة في المعارف والاطلاع. ومثل هذا يتطلب أن يتمتع المحتوى على الإنترنت بجودة تحاكي جودة الكتب المعرفية على الأقل. أي من المفترض أن تكون الجودة أعلى بكثير مما هي عليه اليوم، نتيجة لإمكانية التعاون بين الأفراد والمؤسسات في بناء المعلومات وتوليد المعارف، وكذلك نتيجة لكل وسائل تقديم المعلومة من نص وفيديو وصوت. وهذه الجودة ضرورية لتعزيز ثقة المستخدم بالمحتوى وبالتالي الوصول إلى الفوائد المرتقبة.

تعالج هذه الدراسة موضوع معايير الجودة الضرورية للارتقاء بالمحتوى الرقمي العربي. يقدم الفصل الأول حقائق بخصوص المحتوى العربي بالدرجة الأولى، ويبين الفصل الثاني معايير إنشاء المواقع على الإنترنت الخاصة بالمحتوى نفسه وبالجوانب الفنية لبناء المواقع ومعايير قياس النجاح. أما الفصل الثالث فيستعرض تطور المحتوى على الوب ماضياً وحاضراً ومستقبلاً، ويخصص الفصل الرابع للتوصيات الخاصة بتطوير المحتوى العربي وتحسين جودته.

أولاً - المحتوى الرقمي العربي وأدواته

يرتبط النقاش حول المحتوى الرقمي ضمناً بجودة هذا المحتوى وفائدته. وتصبح هذه الجودة ضرورة ملحة مع تسارع التطور التكنولوجي في مجالي المعلومات والاتصالات وانتشار منتجات هذه التكنولوجيا بسرعة كبيرة، ويدل على ذلك سرعة انتشار الحزمة العريضة والهواتف الذكية التي تسمى بهواتف الجيل الثالث (3G) وهواتف الجيل الرابع (4G)، وكذلك أدوات الاتصال بشبكة الإنترنت في الأماكن العامة أو عن طريق شبكة اتصالات الهاتف الخلوي. كل هذا يقتضي ضرورة ارتقاء المحتوى إلى مستوى آخر يفوق بعشرات المرات ما كان عليه في مطلع الألفية.

ألف- معطيات أولية

١- انتشار الإنترنت والحزمة العريضة

تخطت نسبة انتشار الإنترنت ٤٠% عالمياً. وهذه النسبة متقاربة مع نسبة انتشار الإنترنت في المنطقة العربية، وهي نسبة لا بأس بها إذا أخذت بالحسبان معدلات الأمية[٤] في الدول العربية، وكذلك عدد الدول العربية التي تعيش أوضاع عدم استقرار من العراق والجمهورية العربية السورية واليمن وليبيا، حيث تتراجع الجهود التنموية أمام جهود تأمين الاستقرار.

ويشير تقرير الاتحاد الدولي للاتصالات للعام ٢٠١٤[٥] إلى أن نسبة انتشار الهواتف الذكية قد تضاعفت في العالم بالنسبة إلى ما كانت عليه في عام ٢٠١١، وأن نسبة الزيادة في الدول النامية بلغت تقريباً أربع مرات ما كانت عليه عام ٢٠١١. وقد بلغت نسبة من لديهم هواتف ذكية في العالم ٣٢%، أي ما يعادل ٢,٣ مليون هاتف. هذه الهواتف الذكية هي الأداة الرئيسية للنفاذ إلى الإنترنت والتطبيقات والخدمات الإلكترونية، أي للمحتوى الرقمي. وقد شددت النقاشات مطلع الألفية على أهمية التواصل عبر الإنترنت من أي مكان وفي أي وقت، ولكن تحقيق ذلك كان يتطلب وجود حاسوب واتصال بالإنترنت. أما اليوم فهذا أمر محقق فعلاً بفضل الهواتف الذكية وتطبيقات التواصل المختلفة.

وقد قدر عدد الهواتف الذكية التي وزعت في الأسواق التجارية لعام ٢٠١٤ بأكثر من بليون جهاز[٦]، أي أن سوق أدوات الاتصال، الهاتفي والشبكي، هي سوق متصاعدة بسرعة بحسب إحصاءات الاتحاد الدولي للاتصالات. وكذلك الأمر بالنسبة للحواسيب اللوحية التي انطلقت عام ٢٠١٠ ليبلغ حجم سوقها اليوم أكثر من ٢٠٠ مليون حاسوب تسوّق سنوياً، وبذلك تضاهي سوق الحواسيب المحمولة. وهذه الزيادة في انتشار الحزمة العريضة، وفي الهاتف النقال أو الثابت، تعكس الحاجة المتزايدة للمحتوى بصرف النظر عن طبيعته ونوعه.

---

[٤] تبلغ نسبة من يعرف القراءة والكتابة في العالم العربي ٧٧% لمن تجاوزت أعمارهم ١٥ سنة وفق تقرير التنمية البشرية لعام ٢٠١٤ الصادر عن برنامج الأمم المتحدة الإنمائي.

[٥] ITU. The 2014 ICT Facts and Figures. https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2014-e.pdf

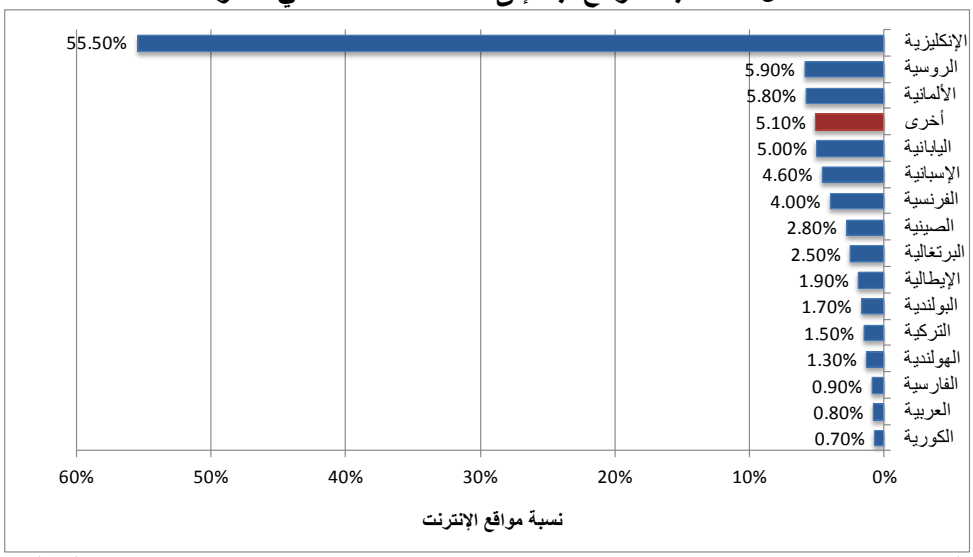
[٦] http://digitalmarketingphilippines.com/amazing-facts-and-statistics-about-visual-web-infographic/

أما في العالم العربي فتقدر نسبة الهواتف الذكية بـ ٢٥% لعام ٢٠١٤ في حين أنها كانت تقدر بـ ١٩% لعام ٢٠١٣، وهي بالتالي في تزايد مضطرد. أما نسبة المشتركين في الحزمة العريضة الثابتة فلا تزيد عن ٨% وهي ليست بعيدة جداً عن الوسطي العالمي المقدرة بـ ١٠%.

٢- حجم المحتوى العربي

تطغى أعداد المواقع التي تضم صفحات مكتوبة باللغة الإنكليزية حتى اليوم على كل المواقع باللغات الأخرى، وذلك بصرف النظر عن منشئها، إن كان حكومياً أو خاصاً أو صادراً عن قطاع الأعمال، كما أن نسبتها أعلى من نصف عدد المواقع مجتمعة لكل اللغات كما يبين الشكل ١[٧].

الشكل ١- نسب المواقع تبعاً إلى اللغات المستعملة في محتواها



المصدر: https://en.wikipedia.org/wiki/Languages_used_on_the_Internet كما يمكن الاطلاع على الموقع التالي: http://w3techs.com/technologies/overview/content_language/all

ملاحظة: بنيت معلومات هذا الشكل على أساسٍ اعتُمد فيه العشرة ملايين موقع الأولى بحسب تصنيف أليكسا[٨] وأهمل الباقي كونه يمثل مواقع لا قيمة لها من حيث عدد زياراتها.

وقد تطرقت الإسكوا في دراسات سابقة إلى إشكالية قياس المحتوى الرقمي العربي، والتي توازي إشكالية قياس المحتوى بشكل عام[٩]. فبعض التقديرات الأخرى، لاتزال تشير إلى تراوح نسبة المحتوى الرقمي العربي على الإنترنت ما بين ٢ و٣%. أما الشكل السابق، فيظهر وضوحاً ضعف

[٧] http://w3techs.com/technologies/overview/content_language/all

[٨] يعتمد هذا التقرير على ترتيب المواقع بحسب أليكسا alexa.com كون الدراسة المشار إليها تعتمد على هذا الموقع من جهة ولكونه أول المواقع التي قامت بعملية ترتيب المواقع، ولأن مواقع الترتيب الأخرى مثل similarweb.com تعتمد المبادئ نفسها ولكن تختلف في مصادر المعلومات. وهناك إجماع عام على مصداقية أليكسا وضخامة قاعدة معطياتها وصغر مدة تحديث بياناتها مقارنة بمواقع الترتيب الأخرى. وفي كل الأحوال لن يكون الخلاف كبيراً في لائحة أولى المواقع العالمية مهما كانت الجهة التي تقوم بالترتيب. أضف إلى ذلك أن موقع أليكسا يعطي خيارات لا تعطيها مواقع الترتيب الأخرى، مثل ترتيب المواقع بحسب المناطق الجغرافية في العالم وكذلك ترتيب المواقع بحسب اللغات وهذا أمر مهم في هذا التقرير.

[٩] لمزيد من المعلومات حول قياس المحتوى الرقمي العربي، يرجى الاطلاع على نسخة عام ٢٠١٥ من تقرير الملامح الإقليمية لمجتمع المعلومات في المنطقة العربية.

حضور المواقع العربية بين مجموع المواقع العالمية من حيث استخدام هذه المواقع، خاصة إذا أخذنا بالحسبان أن متحدثي العربية يمثلون ٥% من سكان العالم وأن ٤,٨% من مستخدمي الإنترنت هم من المنطقة العربية.

باء- قضية المحتوى الرقمي

تعتبر شبكة الإنترنت بالمبدأ مجالاً مفتوحاً للنشر، وهي تمكن المستخدمين من نشر وتحميل كل ما هو رقمي ويمكن تداوله. وهذه الميزة الكبيرة تشكل خطورة كبيرة في الوقت نفسه. فالميزة تكمن في الاستفادة من هذا الفضاء المفتوح في مختلف جوانب الحياة. أما الخطورة فتكمن في انتشار المحتوى الغث، الذي لن يدوم بالطبع أمام محتوى آخر يتمتع بجودة عالية، بشرط توفر هذا الأخير. ومثل هذه الخطورة لا تمثل مشكلة حقيقية في الدول المتقدمة حيث يأخذ الأفراد والمؤسسات أدواراً مبادرة تجعل من المحتوى قضية سياقية تماماً، بمعنى أن هناك من ينتج محتوى وهناك من يُقبل عليه. فهذه المجتمعات قدمت بمبادراتها معظم المنتجات والتطبيقات الواسعة التي تشغل الإنترنت من الموسوعة الحرة ويكيبيديا إلى مواقع التواصل الاجتماعي ومحركات البحث وما تفرّع عنها من منتجات محتوى. أما في المجتمعات النامية حيث تضعف المبادرة وحيث الإمكانيات محدودة ونسبة التعلم العالي منخفضة مقارنة بالدول المتقدمة، وحيث سيطرة الدولة واضحة على معظم المفاصل، فستكون المسؤولية ملقاة على عاتق المؤسسات، وخاصة الحكومية منها أو شبه الحكومية، وعليها الدور الأول في التعاطي مع قضية المحتوى. ومن هنا ضرورة أن تتضمن الاستراتيجيات العربية لتكنولوجيا المعلومات والاتصالات فصلاً خاصاً بالمحتوى الرقمي العربي وما يرتبط به من فرص وتحديات، أو وضع استراتيجيات مخصصة للمحتوى الرقمي. انظر الإطار ١ لبعض الأمثلة حول الاستراتيجيات الرقمية في العالم.

وتجدر الإشارة إلى أن مسألة المحتوى هي مسألة سياقية، تتعلق بالعرض والطلب، أكثر من كونها مسألة يمكن مواجهتها بالتوصيات والمشاريع المحددة، مثال ذلك انتشار شبكات التواصل الاجتماعي في العالم العربي[١٠]. وهو انتشار سببه أولاً الحاجة الكبيرة إلى التعبير الشخصي، نظراً للجو السياسي والاجتماعي غير المحفز للتعبير الفردي والجماعي بوسائله التقليدية. ثانياً، لأن المجتمع العربي هو من أكثر المجتمعات الإنسانية شباباً حيث أن خط العمر الفاصل[١١] هو ٢٤,٦ سنة بحسب تقرير التنمية البشرية لعام ٢٠١٤، مقابل أكثر من ٤٠ سنة في المجتمعات المتقدمة جداً، والشباب هم الأكثر تداولاً للتكنولوجيات الجديدة.

وتظهر آثار العرض والطلب بشكل جليّ في الصحف العربية، المعروفة في المنطقة العربية على الأقل، حيث لغالبيتها موقعٌ على الإنترنت، ويمكن عبره مطالعة معظم مقالاتها، وهي تحقق بسبب هذا المواقع مكاسب مالية عبر الإعلانات مستغلة في ذلك مكانتها في الصحافة الورقية. كذلك فإن أعمال البيع والشراء على الإنترنت لا تزال محدودة الحجم في المنطقة العربية، ومن ثم فإن كل ما يقابل ذلك من محتوى لا يمكن أن يكون إلا منخفضاً ولن يتغير إلا بتغير هذه الأسباب. ومثل هذا سينطبق على الكثير من الجوانب الأخرى المرتبطة بالمحتوى.

---

[١٠] تخطى معدل انتشار الفيسبوك نسبة ٢١%عام ٢٠١٤ في المنطقة العربية ، وذلك بحسب تقرير كلية محمد بن راشد للإدارة الحكومية. http://www.arabsocialmediareport.com

[١١] وهو العمر الذي يفصل المجتمع إلى قسمين متساويين من حيث العدد.

وفي هذه الإطار، لم يعد المحتوى هو الملك، وكذلك لم يعد مستخدمو الإنترنت هم ملوك المحتوى، بل هو السياق العام للبيئة التي تولّد المحتوى وتتلقى ما يفيدها منه. والسؤال الأهم هو: هل يمكن أن تكون قضية المحتوى منفصلة عن القضايا الهامة الأخرى أم هي انعكاس للحالة الراهنة للمجتمع الذي يولّد هذا المحتوى؟ يمكن الوصول إلى جزء كبير من الإجابة بدراسة محتوى مواقع الجامعات ومراكز البحوث والمؤسسات المنتجة للمعلومات والمعرفة العربية على الإنترنت والخدمات التي تقدمها لطلابها ومجتمعاتها، ومقارنة ذلك مع مواقع مؤسسات مماثلة في دول متقدمة أو على طريق التقدم.

الإطار ١- أمثلة لبعض الاستراتيجيات الجديدة لتكنولوجيا المعلومات والاتصالات

تسعى استراتيجيات تكنولوجيا المعلومات والاتصالات في العقد الثاني من الألفية الحالية إلى تعظيم الفائدة من هذه التكنولوجيا التي أصبحت بنيتها الأساسية مستقرة جداً خاصة في الدول المتقدمة. وهذه الاستراتيجيات تتوجه نحو التطبيقات التي تعد محتوى عملي بامتياز.

ماليزيا الرقمية[١٢]**:** وضعت ماليزيا برنامجاً وطنياً يهدف للتحول نحو اقتصاد رقمي متطور بحلول عام ٢٠٢٠. ويتوقع من البرنامج أن يتيح بيئة لتعزيز استخدام تكنولوجيا المعلومات والاتصالات في كل جوانب الاقتصاد لربط المجتمعات المحلية كلها والتفاعل آنياً بما يؤدي إلى زيادة الدخل الوطني الإجمالي، ويحسّن الإنتاجية ومستوى الحياة. وحدد البرنامج ثلاث قوى دافعة استراتيجية لها القدرة على تحقيق أهداف البرنامج، وهي:

- *التزويد الموجه بحسب الطلب*: لقد حققت ماليزيا تقدماً كبيراً في إرساء البنية الأساسية لتكنولوجيا المعلومات والاتصالات. وتكمن حاجتها الآن في تعظيم الاستفادة من هذه البنية الأساسية قبل أن تصبح بالية. لذا سيكون هدف "ماليزيا الرقمية" تطوير الصناعات الجديدة.
- *الاستهلاك الموجه نحو الإنتاج*: يستهلك الماليزيون الكثير من المعلومات ويستخدمون الإنترنت على نحو كبير، وعليهم الآن أن يستخدموا الإنترنت لتوليد العائدات. وعلى "ماليزيا الرقمية" أن تطور منصات وبرامج لكي تساعد مستخدمي الإنترنت أن يكونوا منتجين-مستهلكين.
- *الانتقال من الإضافة المعرفية المنخفضة إلى الإضافة المعرفية العالية*: ذلك أن الشركات الماليزية تستخدم تكنولوجيا المعلومات والاتصالات على نحو ضعيف في أعمالها، وكذلك القليل من المعارف في مخرجاتها. لذا فإن "ماليزيا الرقمية" ستقوم بما يلزم لدعم نماذج الأعمال الرقمية المجددة، التي ستقود عملية توليد منتجات وخدمات عالية القيمة.

إيرلندا: وضعت إيرلندا استراتيجية وطنية رقمية عام ٢٠١٣ بعنوان "استفادة أكبر من الرقمي"[١٣]. وهي استراتيجية محتوى في أساسها كما يتبين من توجهاتها الرئيسية الأربعة:

- *التوجه الأول*: التداول عبر الانترنت وريادة الأعمال التجارية للسكان الأصليين.
- *التوجه الثاني*: مشاركة أكثر للمواطن بقصد خفض عدد "غير المستعملين للإنترنت إلى النصف" بحلول عام ٢٠١٦.
- *التوجه الثالث*: التربية والتعليم الإلكتروني، وذلك بقصد الاستفادة من إمكانات تكنولوجيا المعلومات والاتصالات في التعليم.
- *التوجه الرابع*: التدابير الحكومية البينية واستراتيجية للحكومة الإلكترونية. ذلك أن مستوى التفاعل الذي يظهره القطاع الحكومي في تقديم خدمات الإنترنت والمعلومات هو عنصر هام في دفع الطلب على

[١٢] http://www.digitalmalaysia.my/about-digital-malaysia/overview

[١٣] http://www.dcenr.gov.ie/NR/rdonlyres/54AF1E6E-1A0D-413F-8CEB-2442C03E09BD/0/NationalDigitalStrategyforIreland.pdf

التكنولوجيا الرقمية والمهارات المرافقة.

الاتحاد الأوربي[١٤] الذي وضع عام ٢٠١٠ استراتيجيته الخاصة بتكنولوجيا المعلومات والاتصالات بعنوان "جدول أعمال أوروبا الرقمي" ضمن إطار ما يسمى "أوروبا ٢٠٢٠". وتحدد سبعة مجالات رئيسية هي:

١. إنشاء بيئة تنظيمية للحزمة العريضة جديدة ومستقرة؛
٢. توفير بنية أساسية لخدمة رقمية عامة جديدة من خلال ربط منح القروض أوروبياً؛
٣. إطلاق التحالف الكبير الخاص بالمهارات الرقمية والوظائف؛
٤. اقتراح استراتيجية وتوجيهات للأمن السيبراني للاتحاد الأوروبي؛
٥. تحديث إطار حق المؤلف في الاتحاد الأوروبي؛
٦. تسريع الحوسبة السحابية من خلال القوة الشرائية للقطاع العام؛
٧. إطلاق استراتيجية صناعية جديدة في الالكترونيات: "إيرباص من رقائق".

يمكن للمحتوى العربي أن يتطور عندما تصبح قضية التنمية، بكل أشكالها، هي القضية الأولى. وسيتطور المحتوى العربي جداً عند إدراك المجتمعات وقياداتها بأن العصر الراهن والمستقبل هو عصر المعرفة، وعندما تضع هذه المجتمعات استراتيجيات للتحول إلى مجتمعات معرفة، خاصة وأن اقتصادات هذه المجتمعات تقوم حتى الآن في الأغلب على الثروات الطبيعية التي لا تدوم، والتي لا تماثل إطلاقاً في قيمتها ما يمكن أن تصل إليه هذه المجتمعات عند تحولها إلى طريق مجتمعات المعرفة.

وبانتظار مثل هذا التحول، يقع الدور الأكبر على المؤسسات للتحضير للعمل على إغناء المحتوى الرقمي، وبالأخص المؤسسات الأكاديمية من جامعات ومعاهد ومراكز بحوث والجمعيات العلمية والأدبية ومؤسسات قوامها المعلومات والمعرفة كالمتاحف والمسارح وأقنية التلفاز ودور النشر والمجلات ومجامع اللغة العربية، والمكتبات الوطنية، وغيرها. إضافة إلى ذلك يوجد دور هام لمؤسسات المجتمع المدني، والذي يتمحور حول نشر محتوى رصين موثوق يهدف إلى رفع السوية المعرفية للمجتمعات العربية عموماً. كما أن على الدول إنشاء قواعد بيانات للمعلومات الوطنية وإتاحتها للناس عبر الإنترنت طالما أن في ذلك فائدة.

جيم- المواقع الأولى في العالم العربي من حيث الاستخدام

تبدأ لائحة مواقع المحتوى العربي العشرة الأولى بحسب ترتيب أليكسا[١٥] ، بمحرك البحث غوغل، وموقع الاستفهام عن مواضيع ومعلومات Ask.com، والبريد الإلكتروني gmail وغوغل السعودي والإماراتي. ويأتي الموقع الرياضي "كورة" في الموقع السادس يليه الموقع الإخباري الإلكتروني "اليوم السابع" ومن ثم موقع للألعاب الترفيهية (http://www.goodgamestudios.com) وبعده الموسوعة الحرة العربية ويكيبيديا، وأخيراً موقع شبكة أقنية إم.بي.سي التلفزيونية.

[١٤] http://ec.europa.eu/digital-agenda/digital-agenda-europe

[١٥] http://www.alexa.com/topsites/category;3/World/Arabic

ويلاحظ في المواقع المائة الأولى في العالم العربي[١٦]، وفقاً لترتيب أليكسا، أن مواقع الأخبار هي المواقع الأكثر استخداماً. وتشمل هذه المواقع مواقع إخبارية لصحف إلكترونية لا مقابل ورقي لها، ومواقع أخبار لصحف لها مقابل ورقي. وتشغل مواقع الأقنية التلفزيونية الرقمية، ومواقع الرياضة، ومواقع البيع (Souq.com) أو مواقع الإعلان عن عروض للبيع أو رغبات في الشراء (http://sy.opensooq.com) مركزاً متقدماً من حيث الاستخدام. وكذلك فإن المواقع الخاصة بالسيدات وربات البيوت، والمواقع الترفيهية والدينية والمعاجم اللغوية تحتل موقعاً جيداً. وكذلك المواقع الدولية مثل موقع الأمم المتحدة والبنك الدولي ومنظمة الأغذية والزراعة.

ويلاحظ غياب المواقع الحكومية والخدمات الحكومية من قائمة المواقع الأكثر استخداماً. وقد يكون هذا مفهوماً لأن لكل بلد عربي مواقعه الخاصة به، والتي لا يشترك فيها عادة مع بقية البلدان العربية. ومع ذلك، فلو أخذنا بلداً عربياً مثل مصر لوجدنا النتيجة السابقة نفسها، علماً أن مصر والسعودية هما الدولتان العربيتان الأكثر نفاذاً إلى الإنترنت، من حيث أعداد المستخدمين.

وكذلك الأمر بالنسبة لمواقع المصارف على الإنترنت، فقد أنشأت العديد من المصارف مواقع لها وأتاحت فرص الاستعلام عن الحسابات المصرفية لزبائنها وخدمات أخرى يختلف مدى تطورها من بلد إلى آخر ومن مصرف إلى آخر، ولكن هذه المصارف لا تزال بعيدة في أتمتة عملياتها عن مستوى المصارف في الدول المتقدمة حيث تتوافر أنظمة رقمية متكاملة.

وعند النظر في المواقع الأولى لكل دولة عربية يمكن ملاحظة نوع من التشابه وفقاً لما سبق، ولكن يضاف إليها موقع فيسبوك الذي يأتي على الأغلب في المرتبة الأولى أو الثانية بعد الموقع العام لغوغل والذي يسبقه في بعض الأحيان موقع غوغل نفسه ولكن الخاص بالدولة نفسها مثل غوغل السعودية أو الإمارات أو العراق. ويلي هذه المواقع في المرتبة الثالثة على الأغلب موقع الفيديو يوتيوب وبعدها يأتي موقع ويكيبيديا العربية والأخبار والرياضة.

في هذا فائدة بلا شك لمستخدمي الإنترنت العرب، لكنها فائدة يمكن أن تكون أعم وأشمل وذات انعكاسات اقتصادية واجتماعية تساهم فعلاً في التنمية، وخاصة في مستويات الخدمات والتعليم الإلكتروني شبه الغائبين في الإنترنت العربية ومحتواها. ومثل هذه الصورة سنجدها في الكثير من الدول ولكن بتنويعات تختلف من بلد إلى آخر. ففي دولة مثل فرنسا، تظهر مواقع التسوق الإلكترونية مثل موقع أمازون بترتيب خامس، وكذلك مواقع المصارف المحلية، والمواقع المرتبطة بالخدمات الحكومية. مما يشير إلى أن للإنترنت ومحتواها دوراً أكثر حضوراً وفاعلية في الحياة الاقتصادية/الاجتماعية.

دال- مبادرات المحتوى العربي

كانت الإسكوا من أوائل المنظمات التي نبّهت لمسألة المحتوى الرقمي العربي، وأعدت منذ العام ٢٠٠٣ العديد من الدراسات، كما نظمت الاجتماعات والندوات المتخصصة بغية تعزيز المحتوى

[١٦] المرجع السابق.

الرقمي العربي وتطوير صناعته[١٧]. كما شمل ما قامت به الإسكوا إطلاق مبادرات لتحفيز الشباب على التوجه إلى إحداث شركات ناشئة تعمل في مجال المحتوى الرقمي العربي، وبالتالي قامت بعدة أنشطة خاصة بالتوعية وبناء القدرات وتنظيم المسابقات. هذا بالإضافة إلى تعاونها مع منظمات إقليمية وحاضنات أعمال تكنولوجية في الدول الأعضاء في سبيل العمل على تعزيز المحتوى الرقمي العربي وصناعته.

كما قامت الجامعة العربية باحتضان مجموعة من اللقاءات حول المحتوى العربي على الإنترنت وموّلت بعض الدراسات، ورعت بعض المشاريع التي تصب في موضوع المحتوى العربي، والتي كان منها تأكيد القمة العربية عام ٢٠٠٨ في دمشق على أهمية المحتوى الرقمي العربي. أضف إلى ذلك الأنشطة المتنوعة للاتحاد الدولي للاتصالات في دراسة ودعم المحتوى الرقمي العربي والتأكيد على أهميته في قمة توصيل العالم العربي التي نظمها الاتحاد عام ٢٠١٢.

كما أطلقت الدول العربية العديد من المبادرات الوطنية لتعزيز المحتوى الرقمي العربي، ومنها مبادرة الملك عبد الله للمحتوى العربي التي أطلقتها المملكة العربية السعودية عام ٢٠٠٩[١٨]، وتقوم جامعة الملك عبد العزيز بمتابعة تفعيل هذه المبادرة ورعاية مختلف أنشطتها. كما أنشأت هيئة تنظيم الاتصالات في الإمارات صندوقاً لتكنولوجيا المعلومات والاتصالات وخصّصت فيه مساراً خاصاً للمحتوى الرقمي العربي[١٩]. وقد وضعت قطر برنامجاً للمحتوى الرقمي ضمن الخطة الوطنية للاتصالات وتكنولوجيا المعلومات لعام ٢٠١٥. وتنظم مؤسسة الكويت للتقدم العلمي سنوياً جائزة الكويت للمحتوى الإلكتروني بهدف إثراء المحتوى الإلكتروني المحلي[٢٠]. هذا غير المبادرات الأخرى[٢١]، الفردية أو التي يقف وراءها بعض من الأفراد، والتي لا تزال بحاجة لمزيد من الدراسة لمعرفة أثرها الحقيقي وإمكانية تحقيقها لخرق في المحتوى العربي.

وقد اهتمت مواقع مؤسساتية عديدة بالمحتوى العربي النصي من بينها مكتبة الإسكندرية[٢٢] التي تتيح تحميل أكثر من عشرين ألف كتاب. كما يمكن قراءة الكتب المتضمنة في مكتبتها الإلكترونية، وهي كتب مصورة. وتقدم المكتبة على موقعها منتجات أخرى غير الكتب، مثل المتاحف التابعة للمكتبة، والبانوراما الحضارية عن مصر وتاريخها. إلا أن هذه النوع من المحتوى الإلكتروني كالكتب المصورة، لا يتم حتى الآن قياسه كجزء من المعارف الذي تتضمنها الإنترنت لأنه لا يدخل ضمن قواعد البيانات التي يمكن البحث فيها. وهنا تظهر أهمية أن يكون المحتوى قابلاً للبحث من خلال محركات البحث العامة والمتخصصة أي أن يكون الشكل التقني للمحتوى الرقمي مطابقاً لتقنيات البحث المعتمدة بما في ذلك الكلمات المفتاحية المطلوبة. وهذا على عكس ما قام به موقع آخر

---

[١٧] الإسكوا. وضع صناعة المحتوى الرقمي العربي في المنطقة العربية (باللغة الإنكليزية) ٢٠١٢. E/ESCWA/ICTD/2012/Technical paper.4

[١٨] http://www.econtent.org.sa/AboutInitiative/Pages/AboutInitiative.aspx

[١٩] صندوق قطاع الاتصالات (http://www.ictfund.ae/arabic/Arabic_Digital_Content-A.html) retrieved 8 July 2013.

[٢٠] جائزة الكويت للمحتوى الإلكتروني. http://www.kuwaiteaward.org.kw .

[٢١] http://taghreedat.com/

[٢٢] http://www.bibalex.org/home/default_ar.aspx

يُعرف باسم "الوراق"[٢٣] الذي قام بإعادة تنضيد كتب التراث المنشورة على موقعه نظراً لصعوبة استخدام أدوات التعرف الضوئي للمخطوطات اليدوية.

وقد حاولت المجتمعات الثقافية المختلفة الحفاظ على إرثها الثقافي المدوّن من نصوص ووثائق ومخطوطات وكتب وأفلام سينمائية وأصوات مسجلة، وأتاحتها على الإنترنت بصيغتها الرقمية، بفضل ما سُمي بالرقمنة (digitization).

هاء- أدوات معالجة اللغات الطبيعية في خدمة المحتوى الرقمي

تعمل مجموعة كبيرة من مخابر البحث العلمي في الجامعات والشركات، العربية والأجنبية، على قضايا تخص اللغة العربية ومعالجتها رقمياً، من تعرف النص، وتعرف الكلام، والتركيب الصوتي، والتدقيق اللغوي، والصرف الآلي، والتشكيل الآلي، وغيرها. وتساهم بعض هذه القضايا بشكل كبير في توسيع الاستفادة من محتوى الإنترنت، مثل رقمنة النصوص المطبوعة، والترجمة الآلية، وتحويل النص إلى كلام مسموع، وتحويل الكلام المسموع إلى نص، والتدقيق الإملائي، والتشكيل الآلي الذي تختص به اللغة العربية من بين الكثير من اللغات الأخرى. ولعل أهمها في المرحلة الحالية هي الرقمنة والترجمة الآلية. يرجى الاطلاع على الملحق ١ للائحة بأبرز محاور البحث والتطوير في معالجة الكلام والنصوص في المنطقة العربية.

١- الرقمنة

والمقصود بالرقمنة تحويل النص المكتوب أو المواد المسجلة على أدوات تخزين كانت مستعملة في العقود الماضية، من صوت وأفلام وفيديو، إلى محتوى رقمي. فبترقيم هذه المحتوى، يمكن حفظه بصيغة رقمية وكذلك تداوله عبر الإنترنت وقراءته أو سماعه أو مشاهدته. ولهذا أهمية كبرى خاصة في حالة النصوص التي يمكن البحث فيها عند تحويلها إلى رقمية.

وتظهر التجربة أن إجراءات الرقمنة في حالات التسجيلات الصوتية والأفلام أمراً تقنياً بحتاً ولا يمثل صعوبات خاصة. أما في حالة اللغات فالأمر أعقد، إذ ينبغي التعرف إلى أنواع خطوط الكتابة، وهي في اللغة العربية كثيرة جداً، إضافة إلى أن الطباعة في العقود الماضية لم تكن معيارية، وخاصة من حيث الفراغات بين الكلمات أو مقاطع الكلمات. تضاف إلى ذلك مشكلة الحروف المنفصلة والمتصلة والتشكيل وعلامات التنقيط. وهذه المشكلة هي من أكثر المشاكل تعقيداً أمام بناء أنظمة المعلومات الخاصة بالوثائق الرسمية، مثل السجلات العقارية، المكتوبة بخط اليد، أو التي تعاني من التآكل بسبب قدم بعضها الذي يعود إلى عقود خلت وربما أكثر.

وقد جرت أبحاث كثيرة لحل هذه المسألة، وتوصل البحث إلى ما يعرف بتكنولوجيا "التعرف الضوئي على المحارف" التي تقوم بمسح ضوئي لنص مصور ومن ثم تحاول التعرف إلى الكلمات المشكلة للنص عن طريق التعرف إلى حروف الكلمة ومن ثم تجميعها. وقد أحرزت اللغات المكتوبة بالحرف اللاتيني تقدماً كبيراً في هذا المضمار، وجرت محاولات مماثلة بالنسبة للغة العربية ولكنها لم تحرز التقدم نفسه، ولا تزال نسبة الخطأ لا تقل عن ٢٥% في حالة النصوص بطباعة الخطوط

[٢٣] http://www.alwaraq.net/Core/index.jsp?option=1

المتوفرة على الحواسيب[٢٤]. أما في حالة الأحرف الطباعية التي كانت تستخدم حتى نهاية القرن الماضي فتكون نسبة الأخطاء أكبر وتصل إلى حد ٥٠%[٢٥]. وهذه النسب تتغير وفقاً لنوع الخط ولبرمجيات التعرف. وبعض الأخطاء يمكن تقديرها وتصويبها دونما حاجة إلى العودة للنص الأصلي، وبعضها لا يمكن تصويبه إلا بالعودة إلى النص الأصلي، وهذا يعني ضرورة المراجعة والتدخل والتصحيح، وقد يكون الزمن اللازم لذلك أكبر من الزمن اللازم لإعادة الطباعة من جديد، وهو ما قامت به، كما ذكرنا سابقاً، بعض المواقع العربية (الوراق مثلاً) المهتمة بالتراث خصوصاً حيث لا يمكن لتقنيات التعرف الضوئي أن تكون ذات فائدة كبيرة في التعرف إلى المخطوطات اليدوية التي تفتقد إلى معايير واضحة ومعروفة. وهذه المخطوطات هي الأكثر عدداً ذلك أن المطبعة لم تدخل العالم العربي إلا في القرن التاسع عشر.

٢- الترجمة الآلية

مع توفر المحتوى على الإنترنت في كل اللغات، فإن فائدة أكبر ستتحقق إذا أمكن لكل الناس النفاذ إلى محتوى الإنترنت من كل اللغات. وهذا يتطلب طبعاً توفر ترجمة بينية آلية، أي بدون تدخل الإنسان المترجم، وهذه الترجمة ستكون باستخدام نظام حاسوبي. ومثل هذه الترجمة ستكون ضرورية للغات الفقيرة المحتوى مما سيمكّن أصحاب هذه اللغات من الاستفادة من المحتوى المكتوب باللغات الأخرى. وبما أن المحتوى الأكبر حالياً هو الإنكليزي، فمن الضروري على الأقل توفر الترجمة من الإنكليزية إلى اللغات الأخرى.

وقد طرحت مسألة الترجمة الآلية للبحث العلمي منذ عدة عقود، وخاصة للترجمة بين اللغات الغربية عموماً، نظراً لضرورة ترجمة نصوص كثيرة في مختلف المجالات، الاقتصادية والعلمية، التي أصبح يصعب ترجمتها عن طريق مترجمين أشخاص. وأصبحت هذه المسألة مسألة ملحة مع الإنترنت. وبالفعل أدت هذه الأبحاث والدراسات إلى نتائج جيدة في مجال الترجمة الآلية وخاصة بين بعض اللغات مثل بين الفرنسية والإنكليزية. ومثل هذه الأنظمة بعضها متوفر على الشبكة نفسها ويمكن استخدامها مجاناً. ومن بين هذه الأنظمة النظام الذي يستخدمه محرك البحث غوغل الذي يمكّن مستخدميه من الترجمة بين لغات حيّة كثيرة.

ولكن جودة هذه الترجمات متفاوتة، بين لغات ولغات أخرى، وبين أنظمة ترجمة وأخرى. ومن التجربة التي يمكن القيام بها على منتجات برمجية تجارية[٢٦] وبرمجيات متوفرة على الإنترنت، يمكن أن نلمس بسهولة جودة ودقة في الترجمة بين الفرنسية والإنكليزية مثلاً، وهي ليست بالجودة نفسها بين العربية والإنكليزية.

ومرد ذلك يعود في الواقع إلى أسباب موضوعية تتعلق باللغة العربية، فهي لغة شديدة السياقية من حيث رسم الكلمات، وتتعزز هذه المشكلة خاصة إذا غاب التشكيل. هذه الصعوبة تظهر على نحو أكبر عند الترجمة من العربية إلى اللغات الأخرى. ومع هذا فإن الترجمة إلى العربية ليست دائماً

---

[٢٤] http://www.to-text.net/

[٢٥] نتيجة تجربة على ٢٩ صفحة من القطع الكبير صورت بدقة عالية للمعجم الكبير (مجمع اللغة العربية في القاهرة).

[٢٦] http://www.systransoft.com/

في أحسن حالاتها، مما يستدعي متابعة البحث في هذا المجال والوصول إلى ترجمة بجودة عالية خدمة لطلاب محتوى الإنترنت من العرب[٢٧].

نوع آخر من الترجمة سيكون ضرورياً للاستفادة من الإنترنت هو الترجمة الصوتية. إذا أن النسبة الأعلى من حجم المحتوى هي حالياً المحتوى الفيديوي. وبالرغم من إمكانية الترجمة نصياً لكن الفائدة لن تكون بقدر الترجمة الصوتية المباشرة، من الإنكليزي إلى العربي مثلاً.

ويبين الملحق ١ أبرز محاور البحث في معالجة النصوص والكلام، والجهات العاملة في هذا المجال في العالم العربي.

---

[٢٧] لنأخذ الجملة الإنكليزية البسيطة الآتية: In the fourth and final lecture, we examine the notion of 'objective fact' on which scientific theories are built.

وتعطي الترجمة الآلية:

[في المحاضرة الرابعة والأخيرة، ندرس مفهوم" حقيقة موضوعية "الذي تبنى النظريات العلمية.]

وهي جملة غير مكتملة المعنى. أما الترجمة الفرنسية فتعطي:

Dans la quatrième et dernière conférence, nous examinons la notion de «fait objectif» sur lequel les théories scientifiques sont construits.

وهي جملة مكتملة المعنى، مع أخطاء في تصريف المؤنث يمكن للمدقق الإملائي أن يكتشف بعضها.

ثانياً- معايير المواقع والمحتوى

كانت مواقع الإنترنت عند ظهورها شيئاً جديداً لم يكن معروفاً سابقاً، لذا استدعت أن تكون المواد المحمّلة عليها ذات أفكار جديدة، وقوالب جديدة، وتقدم خدمات جديدة لم تكن متاحة سابقاً. لقد تطلبت الإنترنت التجديد في الأفكار والأساليب، فمواقع البيع على الإنترنت مثلاً، بكل الإضافات التي تحتويها، قدمت ما لم تقدمه وسائل التسوق التقليدية، إذ قدمت هذه المواقع قيمة مضافة على ما هو موجود وليس تكراراً له. ومن جهة ثانية فإن محتوى المواقع الموجودة على الإنترنت هو محتوى مفتوح، يمكن للجميع الاطلاع عليه والاستفادة منه. كما أنه ليس موجهاً إلى مجموعة محددة من الزوار. وهذا على غير ما كان معروفاً سابقاً عند استعمال الكتيبات أو النشرات المطبوعة والتي كانت تستهدف جمهوراً بحد ذاته.

كذلك الأمر فإن المحتوى الموجود على الإنترنت هو محتوى له قيمة استراتيجية واقتصادية واجتماعية. فهو، في بعض الأحيان، يتضمن بيانات تنموية، اقتصادية واجتماعية، لها قيمة في الحياة اليومية وبالتالي لا يمكن العبث فيها، مثل معلومات الحالة الجوية. كما أن مواقع الإنترنت تحتاج إلى التحديث الدوري، ليس في المحتوى فحسب، وإنما أيضاً في وسائل وأدوات الاتصال بها من عتاد وبرمجيات.

يستدعي ما سبق اعتبار الأمور الثلاثة التالية:

١) البنية الأساسية الحاملة لهذا المحتوى والمنتشرة بين الدول ومناطق العالم المختلفة، وهذا الحامل هو حامل دولي تقوم بالوصاية عليه وإدارته وتشغيله شركات خاصة وهيئات شبه دولية، بما في ذلك خطوط النقل وبوابات عبور وحواسيب التخديم المخزنة للمواقع ومحتواها. وهذا جانب استراتيجي وبالتالي تقع على عاتق الدول، أو تجمعاتها الإقليمية، وضع سياسات واضحة إزاءها. وهو أمر له خطورته عند تعلقه بالمسائل الخدمية والاقتصادية إذ أن عدم السيطرة على المحتوى قد يعرض سيادة الدول إلى الخطر عندما يتعلق الأمر ببيانات سيادية، خاصة وأن وجود هذه البيانات على مخدمات خارج سيطرة البلد قد تعرضها للسرقة أو التخريب؛

٢) التطور المستمر الحاصل في وسائل الاتصال بالإنترنت من عتاد وبرمجيات، فأدوات الاتصال اليوم هي غير تلك التي كانت متوفرة منذ عشر سنوات، وستكون أيضاً غير تلك التي ستتوفر بعد عشر سنوات، وتطبيقات المحتوى ستستمر بالتطور. وهذا الجانب هو من اختصاص الباحثين ومراكز البحوث. وهو اليوم في حد ذاته دعامة اقتصادية هامة للكثير من الدول التي أبلت بلاء حسناً في هذا الجانب مثل فنلندا وكوريا الجنوبية وطبعاً الولايات المتحدة التي تحصد الحصة الأكبر؛

٣) المحتوى الكامن في مواقع الوب وجودته، وهو أكثر ما يركز عليه هذا التقرير، إذ يجب أن تقدم المواقع المحتوى الصحيح لمستخدم الإنترنت المستهدف في الوقت الصحيح ضمن إطار يضمن تطوير المحتوى اللازم ونشره وإدارته. وهو في هذا يختلف عما كان عليه المحتوى في النشر الورقي أو النشر بأشكال أخرى.

وتتمثل أولى خطوات إقامة مواقع الإنترنت بالتفكير منهجياً في فكرة الموقع وصقلها، ومناقشة الخيارات المختلفة، وذلك عبر وضع دراسة جدوى. فبداية الموقع ستحدد الكثير من عناصر نجاحه وبقائه. وثاني الخطوات هي بناء استراتيجية خاصة بمتطلبات الموقع حاضراً ومستقبلاً. ويقود ذلك إلى ضرورة اعتماد مبادئ ومعايير تأسيسية لبناء المواقع من حيث الاستعمال والفائدة والأهداف والنشر والإخراج والسلوك.

ألف- المبادئ العامة المؤسسة لمحتوى المواقع

أصدر إئتلاف الشبكة العنكوبتية العالمي (W3C) النسخة الثانية[٢٨](WCAG 2.0) من توجيهات تخص إقامة المواقع على الوب، وهذه التوجيهات تخص مواضيع النفاذ (accessibility) والاستعمالية (usability) والتشغيل البيني (interoperability)، وذلك بناء على المبادئ الأساسية الأربعة الآتية:

١. قابلية المعلومات للإدراك (Perceivable): وذلك في كل ما يخص المحتوى من معلومات ومكونات خاصة بالتعامل مع الموقع، وضرورة أن تكون معروضة للمستخدم بطريقة يمكنه أن يدركها، من حيث إمكانية الاستعاضة عن الأشكال غير النصية بأشكال أخرى يحتاجها الناس، من تكبير لخط الطباعة أو خط بريل أو الصوت أو الرموز أو تبسيط اللغة، أو عرضها بطرق مختلفة، من بصري وسمعي. وكذلك تقديم بدائل عن الإعلام المرتبط بالوقت، أي إمكانية الوصول إلى الوسائط الإعلامية في أي وقت يريده المستخدم؛

٢. قابلية التشغيل (Operable): وخاصة فيما يتعلق بمكونات التعامل مع الموقع بحيث تسمح للمستخدم التجول في الموقع بيسر باستخدام الوسائل التي يتيحها الجهاز المستقبل وخاصة لوحة المفاتيح. وكذلك منح المستخدم ما يكفي من الوقت لقراءة واستعمال المحتوى، وألا يتضمن المحتوى ما يمكن أن يتسبب بأزمة للمستخدم. وكذلك تقديم ما يلزم للمستخدم من أدوات توجيه وتحرك داخل الموقع إلى المحتوى المطلوب؛

٣. قابلية الفهم (Understandable): حيث ينبغي أن تكون المعلومات النصية قابلة للقراءة والفهم من حيث المفردات والمختصرات وألا تستخدم الكلمات أو الرموز غير المألوفة، وأن ترفق المصطلحات بتعريفها بحيث يتمكن من قراءة النص من وصل إلى مرحلة دراسية كالمرحلة الثانوية. وكذلك تصميم الموقع بحيث تبدو صفحاته واضحة وتعمل كما هو متوقع. وكذلك تزويد المستخدم بما يلزم لتحاشي الأخطاء وتصحيحها وخاصة في أعمال المداولات الإلكترونية؛

٤. متانة المحتوى (Robust): ينبغي أن يكون النص متوائماً مع كافة أنواع المستخدمين المحتملين حالياً ومستقبلاً، وبما يتلاءم مع الأدوات التكنولوجية المساعدة التي يستخدمها الأشخاص ذوي الإعاقات أو الحتياجات الخاصة وغيرهم.

ووضع الائتلاف (W3C) قواعد إرشادية خاصة بتطبيق كل مبدأ من المبادئ الأربعة. وفي الواقع فإن هذه المبادئ الأربعة هي مبادئ جامعة تؤكد على أن يكون لجميع الناس الفرصة نفسها في الاستفادة من المحتوى وأنه لا يوجد سوى وب واحدة (one web). فهي تنبه مثلاً إلى أن مستخدمي الإنترنت قد يكونوا أصحّاء أو ذوي احتياجات/إعاقات جسدية خاصة ، إلا أن كلاهما ينبغي أن

[٢٨] http://www.w3.org/TR/WCAG20/

يتمكن من الاستفادة من المحتوى بالقدر نفسه. ذلك أن الأشخاص ذوي الإعاقات هم شريحة لها ثقلها في المجتمع وتتزايد مع التقدم في العمر. فهذه النسبة هي بحدود ١٠% لمن تتراوح أعمارهم بين ١٨ و٣٤ سنة، وتتزايد لتصبح حوالى ٢٧% لمن تتراوح أعمارهم بين ٤٥ و ٦٤ سنة، لتصل إلى أكثر من ٥٠% لمن تزيد أعمارهم عن ٦٥ سنة[٢٩]. أما الإعاقات، فهي إعاقات سمعية أو بصرية بالدرجة الأولى وأغلبها تتمثل في صعوبة القراءة وخاصة لنصوص طويلة نسبياً. وهذه النقطة الأخير لا تخص فقط الأشخاص ذوي الإعاقات وإنما الأصحاء ولكن ممن تلقوا تعليماً متواضعاً وتمثل القراءة بالنسبة لهم صعوبة تمنعهم من الاستفادة من الإنترنت على الوجه الصحيح.

وسيزداد ثقل مثل هذ الإعاقات عند مستخدمي الإنترنت نظراً لزيادة وسطي العمر المتوقع عند الولادة في كثير من المجتمعات. فبحسب تقرير صادر عن الأمم المتحدة في عام ٢٠١٣ بعنوان "تقدم العمر في سكان العالم"[٣٠] فإن نسبة من ستزيد أعمارهم عن الستين عاماً ستبلغ ٢١% في العالم في الفترة بين ٢٠٢٠ و ٢٠٢٥.

كما أن هذه المبادئ تؤكد على إمكانية الاستفادة من محتوى الإنترنت بالنسبة للجميع، بحيث يمكن لمن حاز على قدر معقول من التعليم، التعليم الثانوي مثلاً، من فهم المحتوى والاستفادة منه. وهذه النقطة تشير إلى أن المحتوى ينبغي أن يكون قابلاً للفهم دون الرجوع إلى مصادر أخرى أو أن يكون هناك روابط تساعد على فهم هذا المحتوى. كما يجدر أن يسهل النفاذ إلى هذا المحتوى من قبل من لديهم وسائل اتصال سريعة أو الذين تتوفر لديهم وسائل الاتصال بالإنترنت من الجيل الأول.

وفي العموم، يجدر أن يتم بناء المواقع على مبادئ تختص بالنفاذية إلى الموقع كما هي الحال في المبادئ أعلاه، ومبادئ أو توجهات تخص استعمالية معلومات المواقع وخاصة عندما يتعلق الأمر بالخدمات الإلكترونية، وبالأخص منها الحكومية، حيث يجب أن لا تحتمل المعلومات أي لبس. والاستعمالية عموماً ترتبط بالمدى الذي يمكن لزائر الموقع استخدامه بغية تحقيق أهداف محددة بفاعلية ومردودية مقبولين[٣١].

ومع التطور التكنولوجي، ينبغي أن تدعم بنية الموقع النفاذ إليه من قبل الحواسيب اللوحية والهواتف الذكية، أي ما يسمى بالتشغيل البيني، وهذا الجانب الخاص بالمحتوى هو جانب فني بحت. ولكن وفقاً لمبادئ المحتوى الأربعة السابقة، فينبغي أن يكون الوصول إلى المحتوى ممكناً بصرف النظر عن وسيلة الاتصال بالشبكة. ومن أكثر الوسائط التي تثير مشكلة التوافقية مع الوب هي أجهزة الاتصالات الذكية. أما أسباب هذه المشكلة فهي كثرة المكونات التكنولوجية المستعملة حالياً في تقديم المحتوى مثل: (١) لغات تحرير المحتوى وتنسيقه، (٢) الإخراج الفني للصفحات، (٣) الجيل الذي تنتمي إليه هذه الصفحات إن كانت ساكنة (ستاتيكية) من الجيل الأول، أو ديناميكية تفاعلية من الجيل الثاني)، (٤) محركات البحث، (٥) وسيلة الاتصال من حاسوب أو حاسوب لوحي أو هاتف

---

[٢٩] http://uiaccess.com/understanding.html

[٣٠] World Population Aging, 2013, United Nation.

[٣١] ISO 9241-11: Ergonomic Requirements for Office Work with Visual Display Terminals, Part 11: Guidance on Usability

ذكي، (٦) سرعة الاتصال بالشبكة حيث تظهر مشاكل بطء الشبكة بسبب حجم الذاكرة حيث تُختزن البيانات المؤقتة، ومحدودية المعالج، ولوحة المفاتيح وسعة الشاشة.

وقد قدم التجمع العالمي للشبكة العنكبوتية مجموعة توصيات مستقاة من أفضل الممارسات يؤدي احترامها إلى حل هذه المسألة[٣٢]. إلا أن هذه المسألة آخذة في التلاشي مع التطور التكنولوجي الجاري حالياً، سواء في العتاد أو البرمجيات، إلا ما تفرضه القيود الفيزيائية المتمثلة في حجم الشاشة ولوحة المفاتيح.

إن لمثل هذه المبادئ والخطوط الإرشادية الخاصة بالمبادئ الأربعة أهمية خاصة في التطبيقات الحكومية ومن المهم أن تضع الدول مثل هذه الخطوط الإرشادية التي تحدد فيها رؤيتها لكيفية التعامل مع المحتوى، والحدود الفنية والقانونية، ومسؤولية مختلف الجهات المعنية في تنفيذ خدمة إلكترونية ما. وفي هذا الصدد يمكن الاطلاع على التجربة الكندية على الرابط: http://www.tbs-sct.gc.ca/ws-nw/index-eng.asp التي تحدد فيها رؤيتها، المطابقة للقواعد العامة المذكورة سابقاً. ويتضمن الموقع عدداً من المعايير الوطنية كما يبين آليات التنفيذ الإدارية والفنية.

باء- معايير جودة المحتوى

من أجل تحقيق غاية معينة من المواقع على الإنترنت يجب العمل على تحديث محتواها وتطوير التقنيات المستخدمة باستمرار وذلك من أجل تلبية حاجات المستخدمين ومتطلباتهم المتزايدة.

١- معايير المحتوى

على المواقع أن تلتزم بمعايير تخص المحتوى كما:

- **مرجعية الموقع**: سواء كانت هذه المرجعية شخصاً أو مؤسسة مسؤولة عن الموقع، من الضروري أن تكون مؤهلة وقادرة على القيام بإدارة الموقع وتغذيته بالمحتوى اللازم. وذلك مع الحرص على تسمية هذه المرجعية بشكل واضح على الموقع والتعريف بها وبإمكاناتها ونشر عنوانين الاتصال بها، وكذلك تبيان كل الجهات المساندة لهذا الموقع؛
- **القصد من الموقع**: بتوضيح الهدف من المعلومات التي يقدمها الموقع. وهنا ينبغي التحقق من أن يقدم الموقع معلومات تتوافق مع أهدافه، وأن يكون موجهاً فعلاً للمجموعات المستهدفة، وأن تكون الوصلات التي يقترحها الموقع تخدم أهدافه؛
- **الشمولية**: يجب أن يتحقق القائمون على الموقع من انتقاء وشمولية محتوى الموقع، أو وبأن المواضيع المتوفرة معالجة بشكل وافي. كما يجب التحقق من فعالية الوصلات (links) سواء مع المواقع خارجية أم مع ال وصلات داخل الموقع نفسه. وإذا كانت الوصلات مع مواقع خارجية فيجب التحقق من كفاية معلومات الوصلات؛
- **حداثة معلومات الموقع**: وإيقاع تحديثها إذ ينبغي توضيح تاريخ تحرير ونشر المحتوى وكذلك تاريخ آخر تحديث والتحقق من وجود الوصلات وعدم اختفائها؛

[٣٢] http://www.w3.org/TR/mobile-bp/summary

- **الموضوعية**: بحيث تكون معلومات الموقع غير منحازة ولا تتلاعب بزوار الموقع لتحقيق أهداف غير معلنة. وأن تُخضع المعلومات لقواعد المنطق المعروفة التي تربط الأثر بالسبب؛
- **الدقة**: وهو ما تضمنه الجهة الناشرة للمعلومات ومصداقيتها في مجال محتوى الموقع، وكذلك يجب توضيح المراجع المستخدمة كمصدر للمعلومات، بحيث تتوافق هذه المعلومات مع ما هو معروف أو منشور في مصادر أخرى؛
- **اللغة**: من الأهمية بمكان العناية باللغة إذ أن ذلك سيترك انطباعاً حول مدى مهنية الموقع وجدية المشرفين عليه، ولهذا صلة بموثوقية محتواه. ومن المعايير التي يمكن اتباعها في حالة المواقع العربية:
  - احترام القواعد اللغوية الأساسية من حيث الفعل والفاعل والمفعول به، والمبتدأ والخبر، والمضاف والمضاف إليه والقواعد الشائعة مثل العد والمعدود. ومن الأفضل بالطبع احترام كل القواعد إن أمكن وهو أمر غير متاح عموماً إلا للمختصين. ومن الأفضل في هذا الخصوص، وفي حالة المواقع التي تمثل مؤسسات معروفة، الاستعانة بمدقق لغوي تجنباً للالتباسات التي يمكن أن تقع بسبب عدم احترام القواعد اللغوية، تماماً أو الاستعانة بمحرر مختص لتحقيق وضوح المعاني؛
  - استهداف القارئ العادي في الكتابة من حيث وضوح الدلالة وسهولة المفردات وتجنب المصطلحات المتخصصة واستعمال المعرّب الشائع؛
  - تجنب الحشو والاستطراد والتوجه نحو المعلومة المراد إيصالها بشكل مباشر؛
  - تحاشي العرض الطويل المسهب، إذ أن لمستخدم الإنترنت خيارات كثيرة للتصفح وهو سيسعى إلى الوصول إلى ما يريد بأقل زمن ممكن.

وفي هذا الإطار، يجدر بالدول، باعتبارها القائمة على التعليم العام، أن توجه في مناهجها عناية خاصة نحو كتابة التقارير من حيث المضمون والشكل واللغة السليمة. وهذا الأخير أمر هام يمكن ملاحظته بمجرد النظر في الكثير مما هو منشور على الإنترنت باللغة العربية وبين ما هو منشور باللغة الإنكليزية مثلاً. وهذا أمر سيغير من جودة المحتوى العربي بشكل مباشر.

الإطار ٢- العربية العامية والإنترنت

تستخدم العامية في الإنترنت وخاصة على مواقع التواصل الاجتماعي. ويعتبر البعض ذلك تهديداً للغة العربية من حيث أنه يؤكد الكتابة العامية. ولكن هذه الكتابة ليست إلا كتابة مؤقتة تزول بسرعة ولا تمثل محتوى يمكن البحث فيه. وهذا الحال هو جزء من "الخصام" القائم بين الفصحى والعامية وتجاهل أمر أن العامية لم تنشأ من فراغ وإنما من حاجة إلى التعبير السريع البسيط، وهو الدور الأول للغة. صحيح أن العامية لا تملك المفردات التي تملكها الفصحى، وأن قدرة اللغة الفصحى التعبيرية في بعض الجوانب أكبر بكثير من قدرة العامية وخاصة في المواضيع التخصصية، إلا أن من يمتلكونها قلة لدرجة أنه لا يُعرف كتاب، إلا ما ندر، بلا أخطاء في القواعد أو النحو أو الدلالة. كما أن العامية امتلكت من المفردات والتراكيب البسيطة التي تجعل الجميع يتجه نحو استخدامها لا إرادياً.

وبرأينا، كان على الفصحى أن تتطور نحو تسهيل القواعد[٣٣]، وهذا ما لم يحدث. مما أدى إلى أن الدروس تلقى حتى في المدارس ومعاهد التعليم بالعامية. والواقع أن الإنترنت لم تقم بتشجيع الناس على العامية بأكثر مما يفعلونه في تواصلهم اليومي الصوتي، ولا بأكثر من أفلام السينما والمسلسلات التلفزيونية العامية بطبيعتها التي تذهب حد اللهجات المحلية في البلد الواحد، وكذلك الأغاني والشعر العامي المكتوب.

وبالتالي يمكن استغلال الإنترنت للسير نحو لغة وسط، أكثر إمكانية على التعبير. وهذا يتطلب محتوى عربياً ضخماً يجعل من الإنترنت بالنسبة لمتكلمي العربية مصدرهم الأساسي في المعلومات والمعرفة والتواصل.

٢- معايير تكنولوجية للموقع

تعتبر المعايير التكنولوجية جزءاً هاماً من معايير جودة الموقع التي يمكن إيجازها بالآتي:

- **سهولة التجول** في الموقع وسهولة الوصول إلى معلوماته والعودة من صفحة إلى أخرى وفاعلية أداة البحث الداخلي، وعمل وصلاته الداخلية، ووضوح هذه الوصلات بما يسهل على الزائر التنقل في أرجاء الموقع؛
- **سرعة تحميل الموقع** وسهولة الوصول إليه من مواقع أخرى. ووجود خدمة تسمح بإعلام متابعي الموقع بأخبار الموقع أو بعض جوانبه (RSS)، والإشارة إلى كل قسم جديد بما يثير انتباه الزائر مباشرة؛
- **سهولة العثور على الموقع**، عبر نشره لأكبر عدد من الكلمات المفتاحية التي تتيح لمحركات البحث الإنباء عنه والوصول إليه، ومختلف أنواع بيانات البيانات (metadata) التي تسمح بتصنيفه بين المواقع، وتوصيف معلوماته في بعض الحالات باعتباره مصدراً لهذه المعلومات وليس مجرد وثيقة، وكذلك الترويج للموقع عن طريق ربطه بمواقع أخرى؛
- **التوافق مع مختلف محركات البحث**، الخاصة بالحواسيب العادية واللوحية والهواتف الذكية؛
- **المشاركة**، عبر أدوات تسمح للزائر مشاركة الموقع مع آخرين، وإفساح إمكانية متابعي الموقع من التعليق وإبداء الرأي، ومتابعة المواقع لآراء زواره والاستجابة لطلباتهم؛
- **وجود معلومات إحصائية** خاصة بالموقع تسمح لإدارته بقياس نجاحه وإخفاقاته؛
- **الجوانب الفنية** المتعلقة بأمان الدخول إلى الموقع وبيانات مستعمله وكذلك مسائل الملكية الفكرية والخصوصية؛
- **موثوقية الجهة الحاضنة** وما تقدمه من خدمات إدارة وصيانة وحماية من هجمات التخريب، وكذلك عرض الحزمة الذي تقدمه بوابات مخدماتها وسرعة الاستجابة في التفاعل مع متغيرات الموقع، ووجود نسخة احتياطية وغير ذلك من الجوانب الفنية.

[٣٣] انظر في هذا الخصوص كتاب "تبسيط قواعد اللغة العربية على أسس جديدة"، الدكتور أنيس فريحة، بيروت ١٩٥٩. كذلك مجلة الفكر العربي العدد ٦١، أيلول/سبتمبر ١٩٩٠، عدد خاص باللغة العربية وإشكالات المعاصرة.

٣- معايير الإخراج الفني للمواقع

لتصميم الموقع دور أساسي في اجتذاب الزوار إليه، فهو في إحدى جوانبه انعكاس لصدقية الموقع ومهنيته، إضافة إلى الجوانب الفنية التي تثير لدى الزائر رغبة البقاء في الموقع والتجول فيه. ولكل موضوع خصوصيته في تصميم الموقع، فمواقع البيع والتجارة لا تماثل مواقع الجامعات مثلاً، ليس فقط من حيث المحتوى وإنما من حيث الطبيعة والزوار المستهدفين. ومع هذا تبقى هناك جوانب مشتركة يستحسن مراعاتها وفقاً لتجارب مصممي المواقع[٣٤]. ومن أهم الجوانب نجد:

- **اسم الموقع**: الذي لاختياره أهمية بالغة من حيث سهولة تذكره وكتابته بما يسهّل العودة إليه أو إعلام الآخرين به. والمقصود هنا باسم الشهرة المتداول إضافة إلى اسم النطاق. إن لاسم الموقع أهمية كبيرة في اجتذاب مستخدمين، تحت ما يسمى بالثواني الأولى الحاسمة في البقاء وتصفح الموقع أو مغادرته؛
- **بنية الموقع**: حيث يتم التأكيد على الأقسام الرئيسية للموقع والعمل على أن تكون الصفحة الرئيسية واضحة بمحتوياتها للزوار وبحيث تبرز الغرض من الموقع دون أن تكون هذه الصفحة مثقلة بالمعلومات. ومن المفضل أن تتراوح كلمات هذه الصفحة بين ٣٠٠ و ٥٠٠ كلمة. ويفضل أن تكون النصوص بشكل عام موجزة مع إمكانية التوسع فيها في صفحة أخرى. ويُفضل تحاشي المنزلقة الجانبية (scrolling) في الصفحة الرئيسية. وكذلك إمكانية رؤية الصفحة الرئيسية بالكامل بصرف النظر عن برمجيات التصفّح؛
- **معلومات الموقع**: ضرورة وضع تفاصيل الاتصال بالموقع والمسؤولين عنه. ومن الأفضل أن يوضع أكبر قدر من تفاصيل الاتصال، وإمكانية أن يترك الزوار رسائلهم عبر الموقع مباشرة؛
- **المظهر**: على الموقع أن يُبنى بتناغم وتجانس بين الألوان والنصوص كما يلي:
  - انتقاء لونين أو ثلاثة للموقع والإبقاء عليها في كل صفحات الموقع وتجنب الألوان المنفّرة أو التي لا تسمح بقراءة سهلة، والتقليل من رايات الصفحات (banners) ؛
  - تحاشي استخدام الدفق الضوئي أو الإيحاءات المتحركة (animation) والرسوم البيانية التي قد تُبعد الزائر عن المحتوى المراد إيصاله، وإنما استخدامها عند الضرورة مع ملاحظة أن تحميلها قد يستغرق وقتاً الأمر الذي يدفع بالزائر إلى ترك الموقع؛
  - ترك فراغ أبيض في الصفحة ذلك أن الصفحة المكتظة تُجهد الزائر والفراغ في الصفحات يمكّن الزائر من التركيز على المهم.
- **الإشارة إلى المعلومات الهامة** بشكل واضح لجذب انتباه الزائر؛
- **اختيار بنط واضح معروف**، ويمكن استخدام علامات التعداد لعرض المحتوى بطريقة مباشرة وسهلة القراءة؛

[٣٤] http://www.bruceclay.com/design/web_crit.htm

- **تجنب العبارات المنفرّة،** وعلى نحو قاطع، عبارة "المحتوى قيد الإعداد" لما تثيره هذه العبارة في نفس الزائر من إحباط وتعرض ثقته بالموقع للاهتزاز، ومن الأفضل في هذه الحالة عدم وجود الرابط؛
- **وضوح العناوين** حيث ينبغي أن تكون عنوانين الروابط في الموقع واضحة لا تحتمل التأويل أو سوء الفهم. ووضع ما يُسهّل على الزائر معرفة محتويات الصفحات قبل الذهاب إليها؛
- **التأكد من صلاحية وعمل الروابط،** الداخلية والخارجية. وهذه الأخيرة تحتاج إلى فحص دوري لإمكانية اختفائها مع الزمن؛
- **خارطة الموقع،** وهي ضرورية لكل موقع مهني لأنها تساعد الزائر على الوصول إلى المعلومات التي يريدها دونما خطأ.

هناك معايير خاصة بالجوانب التكنولوجية وأدواتها يفصلها إئتلاف الشبكة العنكبوتية العالمي[٣٥] بخصوص أدوات التحرير والعرض من نصوص وصور وفيديو وإمكانية استخدامها من قبل كل اللغات والأدوات وخاصة الهواتف الذكية.

٤- معايير أخلاقية

إضافة إلى كل ما تقدم، ينوه البعض بالمعايير الأخلاقية الخاصة بالمواقع ومحتواها، وليس القصد هنا نوعية المحتوى، وإنما بمعنى احترامه لمجموعة القواعد السلوكية الآتية:

- **اتباع قوانين الملكية الفكرية** كما تنص على ذلك القوانين العالمية والمحلية، من حيث منع النسخ والنشر إلا بإذن، والإشارة إلى المصادر بوضوح؛
- **احترام حقوق الإنسان** بحيث لا يتضمن المحتوى أي إساءة متعمدة لفرد أو جماعة أو أثنية أو طائفة أو شعب. كما يجب أن لا يستخدم الموقع للتحرش والاعتداء والإيقاع بالآخرين وخاصة القاصرين؛
- **احترام الخصوصية** لزوار الموقع بحيث لا ينتج أي أذى عن زيارة الموقع واحترام خصوصية المستخدم ومعطياته الشخصية وعدم استعمالها فيما يخرق خصوصية صاحبها أو تداولها مع أي جهة أخرى ولأي سبب كان؛
- **الوضوح في المحتوى،** وخاصة المحتوى المرتبط بمداولات بين الزائر وخدمات الموقع، إذ يجب أن تكون التبادلات واضحة ولا تحتمل التأويل أو سوء التفسير.

ويمكن الرجوع إلى بعض جوانب التشريعات السيبرانية في هذا الخصوص التي تناقش جوانب مثل الملكية الفكرية والخصوصية وغيرها مما ينبغي اتباعه في تشغيل المواقع[٣٦]. ومن المفيد في بعض الأحيان اعتماد الموقع، أو المؤسسة التي تقف خلفه، لمدونة أخلاقية واضحة تحدد علاقة الموقع بزواره وكل ما يمكن أن ينجم عن مثل هذه الزيارة.

---

[٣٥] http://www.w3.org/standards/webdesign

[٣٦] http://css.escwa.org.lb/ictd/1785/d22.pdf

جيم- معايير قياس نجاح الموقع

بما أن المواقع تختلف فيما بينها من حيث المحتوى والهدف منها فإن وجود معايير يمكن تطبيقها على كل المواقع لقياس نجاحها سيكون صعباً. فمواقع البيع مثلاً سيكون معيار نجاحها هو عائدات البيع التي تتحقق عن طريق الموقع، ومواقع الأخبار سيكون عدد النقرات (hits) وعدد المشاركات التي تحدث لمقالاتها على صفحات مواقع التواصل الاجتماعي مثلاً. وفي كل الأحوال فإن مقاييس النجاح أو مؤشراته لا يتم تطبيقها منذ اليوم الأول وإنما هناك فترة انتظار ترتبط بنوع المحتوى.

ومن بين معايير قياس النجاح نجد الآتي:

- **عدد الزيارات**: من أكثر معايير قياس نجاح موقع ما هو عدد الزيارات التي يقوم بها مستخدمو الإنترنت لهذا الموقع. وقد استبدل غوغل (في خدمته التحليلية للمواقع Google Analytics) هذا المعيار بمعيار مكافئ يعكس عدد المرات التي يجري فيها التفاعل مع الموقع. ولكن هذا المعيار غير كاف ولا يعبر بالضرورة عن مدى انتشار الموقع، إذ قد تكون مثل هذه الزيارات من فعل برمجيات أخرى (ربوت) تبحث عن معلومات أو تحاول نشر رسائل على الموقع، أو قد تكون نتيجة البحث عن مواضيع ما وتطابق في الكلمات المفتاحية ومن ثم فإن الزيارة ستكون مجرد زيارة عابرة؛
- **مدة الزيارة**: يشير وسطي مدة الزيارة إلى اهتمام الزوار بالمحتوى. ولكن هذا المقياس أيضاً لوحده قد لا يكون معبراً عن قيمة هذا المحتوى ونجاح انتشاره. فمن يبحث عن معلومة ما سيغادر الموقع فور عثوره على هذه المعلومة، فإن كان الموقع مبني بحيث يكون الوصول إلى المحتوى المطلوب سريعاً، كما هو الحال في المعاجم والموسوعات، فسينخفض وسطي زمن الزيارة، أو قد يزيد في حال كون الوصول إلى المحتوى المطلوب قد يستغرق وقتاً، من الموقع أو من شبكة الاتصال؛
- **تكرار الزيارة**: يشير معيار عدد مرات تكرار الزيارة إلى أن الزوار يجدون في محتوى الموقع ما يبحثون عنه في أكثر من جانب. إلا هذا التكرار قد يكون موسمياً بالنسبة للبعض أو متباعداً مما لا يسمح باعتماد التكرار مؤشراً لقياس النجاح؛
- **الارتداد** Bouncing: وهو يمثل عدد الزيارات (الجلسات) التي تزار فيها صفحة واحدة فقط من الموقع ثم يُغادر الموقع. وهو معيار يشير انخفاضه إلى جودة الموقع؛
- **وسطي عدد الصفحات المشاهدة** في زيارة الموقع الواحدة، التي تعبر عن وجود محتوى يثير فضول الزائر واهتمامه، وهو المكافئ السلبي للمعيار السابق.

قد لا يكفي اعتماد معيار واحد من المعايير السابقة كمؤشر للنجاح، لذا سيكون من الأفضل تشكيل معيار مركب بتوزين يناسب الموقع وطبيعته، وعلى الجهة صاحبة الموقع أن تقيس نجاحها وفقاً لهذه الطبيعة. فمحركات البحث سيكون عدد الزيارات هو مقياس نجاحها، أما الموسوعات فسيكون عدد الزوار ومدة بقائهم في الموقع. أما المواقع الإخبارية فسيكون عدد الزوار ومدة زياراتهم وعدد الصفحات التي يدخل إليها الزائر نفسه، وتكرار الزيارة ومشاركة مقالاتهم على صفحات التواصل الاجتماعي. وسيكون الأمر نفسه بالنسبة للمدونات على أن يضاف إليها حجم تفاعل الزوار مع

محتوى الموقع عبر وسطي عدد التعليقات. أما في مواقع البيع والشراء فسيكون معدل التحويل الذي يترجم عدد الزيارات إلى عدد المشترين باستخدام الموقع إضافة إلى مؤشر الإخلاص الذي يشير إلى معدل العودة إلى الموقع للشراء مرة أخرى وعدد الزوار الجدد.

دال- عينة من مواقع عربية ومعايير الجودة

١- النسخة العربية من الموسوعة الحرة ويكيبيديا

انطلقت الموسوعة الحرة ويكيبيديا (Wikipedia) في مطلع العام ٢٠٠١ على أساس رئيس هو عدم وجود هيئة تحرير مركزية، الأمر الذي خالف الموسوعات الرقمية المعروفة مثل إنكارتا (Encarta) أو بريتانيكا (Britannica). ولم يكن من المخطط في بداياتها أن تكون ما هي عليه الآن. إلا أن نشاط من ساهم في إطلاقها والمواضيع التي كتبت لتكون مرافقة لمواضيع لموسوعة نوبيديا Nupedia وانتشار الفكرة في العالم وتبنيها من قبل عدة لغات جعلها تنفصل عن كل المشاريع التي كان من المزمع أن تشارك فيها.

بلغ مجموع المواضيع التي تحتويها موسوعة "ويكيبيديا" حتى نهاية تشرين الثاني/نوفمبر ٢٠١٤ أكثر ٣٤ مليون موضوعاً كتبت بـ ٢٨٨ لغة مختلفة شارك في كتابتها أكثر من ٥٠ مليون مساهم معروف وكثيرون غير معروفين. وهي تحتل اليوم المرتبة السادسة بين المواقع الأكثر زيارة بحسب تصنيف أليكسا.

ومن بين اللغات التي تتوفر فيها موسوعة "ويكيبيديا" اللغة العربية التي يكتب موضوعاتها أشخاص عرب أو يتقنون العربية. وقد زاد عدد المواضيع المتضمنة في الموسوعة الحرة العربية عن ٣٤٠ ألفاً ساهم في تحريرها قرابة ٩٠٠ ألف شخص من بينهم ٣٥٠٠ مساهم نشط (بمساهمة تحريرية واحدة على الأقل) وبمعدل تحديث جيد مقارنة بكثير من اللغات حيث بلغ ١٩٤ تحديثاً وسطياً [٣٧] في حين أن مثل هذا المعدل هو ٨٨٢ بالنسبة للغة الإنكليزية و ١٩١ للغة الفرنسية ٨٩ للغة الألمانية. وهذا المعدل هو مؤشر جودة في الواقع كون التحديث يتضمن الإضافة وكذلك التصحيح.

أما أكثر اللغات زيارة فهي الإنكليزية، تليها الإسبانية، وتأتي العربية في المرتبة العاشرة بعد الروسية والصينية، وهو ترتيب جيد يعكس اهتمام المواطن العربي بالمعلومة بالرغم من أن ترتيب موسوعة ويكيبيديا العربية من حيث عدد المقالات هو ٢٢ عالمياً وتسبقها مباشرة الإندونيسية والتي تسبقها الفنلندية والنرويجية والفارسية. وهي تحتل المكانة التاسعة بين المواقع التي يزورها المستخدمون العرب للإنترنت وذلك بحسب ترتيب ألكسا للمواقع التي يزورها العرب.

ومن مراجعة لبعض العينات من مواضيعها، يلاحظ أن بعضها مكتوب بلغة واضحة وجيدة وتتناول الموضوع بشكل متكامل، وبعضها الآخر لا يتضمن إلا بعض كلمات، يُؤمل أن تُغنى في المستقبل، وبعضها يتضمن أخطاء يؤمل أيضاً أن تصحح في المستقبل.

[٣٧] https://meta.wikimedia.org/wiki/List_of_Wikipedias

ومن معايير الجودة الخاصة بويكيبيديا نجد معيار عدد الكلمات المستخدم في تناول موضوع ما وكذلك المراجع الملحقة في هذا الموضوع والتي تشير إلى مصداقية المعلومات الواردة فيه. والعينة التي جرى مراجعتها تظهر تحقق الجانب الأول أما الجانب الخاص بتنوع المراجع فهو لم يكن محققاً في الغالب في هذه العينة وذلك على النقيض من النسخة الإنكليزية. كما يحتاج الكثير من مواضيعها للإغناء مقارنة مع ما هو موجود بالإنكليزية كما في حالة موضوع "الرقمنة" مثلاً.

أما من حيث الجوانب الفنية فهي تماثل شبيهاتها في اللغات الأخرى، ولا معلومات خاصة عنها من حيث الارتداد وزمن المكوث ووسطي عدد الصفحات لكل زيارة، إذ إن الإحصائيات في هذه الحالة تتعلق بالموقع العام الذي يضم كل اللغات.

وتحتاج مثل هذه الموسوعة إلى اهتمام خاص وجهود حثيثة لتحفيز المشاركة والمساهمة في تحريرها وخاصة من قبل المختصين ممن يعملون في المجال المعرفي وذلك لأمرين أساسيين: الأول ضرورة وجودة موسوعة عربية موثوقة على الإنترنت تكون مرجعاً لكل من يبحث عن معلومة ومعرفة، والثاني هو أنه، وعلى عكس الكثير من اللغات، لا تتوفر، ولم تتوفر، موسوعة متكاملة عربية يمكن الرجوع إليها كما كان الحال في اللغتين الفرنسية (Universalis) والإنكليزية (Britannica). وقد وعدت "مبادرة الملك عبد الله" العناية بهذه الموسوعة.

٢- <u>موقع اليوم السابع</u>

وهو من أفضل المواقع العربية على الإنترنت، إذ يحتل المرتبة ٣٠٠ عالمياً، والمرتبة السابعة عربياً والمرتبة الخامسة مصرياً. وهو موقع إخباري إلكتروني فاق في ترتيبه كل مواقع الجرائد الإخبارية حتى تلك المشهورة تاريخياً. تتوزع أخباره بين المصرية والعالمية، وأخباره شاملة تغطي مختلف جوانب الحياة اليومية.

بحسب معايير جودة المواقع فهو يتطابق معها على نحو كبير. فهدفه واضح ومعلن، كما أن الجهة التي تمتلكه معروفة ورسالتها واضحة بحيث يتضح مباشرة اسم رئيس التحرير الصحفي. أخبار الموقع موجزة وشاملة، ومصاغة بلغة سليمة وواضحة ويترك فسحة للتفاعل مع القارئ. تظهر فيه الحيادية وهذا ما يرفع من مصداقية المحتوى. يتميز الموقع بسرعة تحميل المحتوى، والوصول السريع إلى موضوع ما على الصفحة. وكذلك فإن الإخراج الفني مريح للزائر ويسهل وصوله إلى غايته من زيارة الموقع دون أن ينشغل بالإعلانات.

الارتداد في زيارة الموقع منخفض نسبياً ٢٤%، وعدد الصفحات التي يذهب إليها الزائر حوالى ٥ صفحات وسطياً ومدة مكوث الزائر على الموقع هي أكثر من ٩ دقائق وسطياً. وهذه المؤشرات أكثر من جيدة عند مقارنتها بموقع لصحفية فرنسية عريقة مثل "اللوموند"، حيث يبلغ معدل الارتداد حوالى ٥٠%، ووسطي الصفحات لكل قارئ ٢,٧ صفحة ووسطي زمن المكوث حوالى ٥,٢ دقيقة. أما الترتيب العالمي لهذه الصحيفة فهو ٦٥٨ عالمياً وتحتل المرتبة ٢٤ في ترتيب المواقع في فرنسا.

٣- <u>موقع camelcamelcamel.com</u>

وعلى سبيل المقارنة يمكن أن نأخذ إحدى المواقع العالمية التي صنفتها مجلة "التايم"[٣٨] من بين أفضل خمسين موقعاً لعام ٢٠١٤[٣٩] وهو موقع يهتم بقضايا التجارة كما يشير إلى ذلك اسمه. والهدف من الموقع هو ملاحقة تغير الأسعار لدى محلات "أمازون" المنتشرة في الغرب حيث يجري الشراء منها عبر الإنترنت. وهذه المحلات أصبحت أحد أكبر محلات الشراء في الدول المتقدمة (أمريكا، كندا، أوروبا، كوريا، وغيرها). يقوم الموقع المذكور بملاحقة التغير في أسعار المعروضة على "أمازون"، ثم يقوم الموقع يومياً بإعلام المشترك فيه بالتغير بسعر سلعة ما، يريد المشترك شراءها، وذلك عبر البريد الإلكتروني أو عبر التويتر. كما يزود المشترك بتغيرات سعر السلعة لفترة ثمانية عشر شهراً خلت بحيث يقرر المشترك في هذا الموقع الوقت المناسب لشراء السلعة التي يهتم بها.

بدأ الموقع نشاطه في عام ٢٠٠٨ وهو يحتل اليوم المرتبة ٢٨٠٠ تقريباً ومرتبته في الولايات المتحدة بحدود ٨٠٠. وهو موقع يستجيب مباشرة للهدف منه. لذا استنسخ في عدة بلدان حيث محلات أمازون مزدهرة مثل الولايات المتحدة وبريطانيا وكندا والصين وفرنسا. وهو متوفر بخمس لغات: الإنكليزية والفرنسية والألمانية والإيطالية والإسبانية.

يمكن التجول في الموقع بسهولة والوصول إلى المبتغى دون أية صعوبة. وصلاته الداخلية سريعة وكذلك فتح الصفحة الرئيسية كما أنه لا يتضمن إعلانات تصرف الزائر عن هدفه أو تثقله بما لا يريد ويشير تنظيمه إلى الحرفية العالية. وتتميز متابعته لأسعار "أمازون" بمصداقية متناهية، وهو يقوم بدور الوسيط ويتقاضى عمولة من المبيعات التي تتم عن طريقه. ولا توجد أية وصلة تنتهي برسالة "قيد الإعداد". وتشرف عليه شركة لها سجلها التجاري وأسماء من يدير الموقع معروفة ويمكن مقاضاتهم عند اللزوم.

وفي مقابل هذه المواقع الجيدة، هناك الكثير من المواقع التي لا تحقق الكثير من معايير الجودة، بل هي مواقع سيئة ومسيئة، ومثل هذه المواقع لن يكون لها دور يذكر في حال ازداد عدد المواقع ذات المحتوى الجيد.

هاء- دراسات في المحتوى الرقمي العربي

صدرت العديد من الدراسات عن المحتوى الرقمي العربي وعن صناعته عن منظمات إقليمية ودولية منها الإسكوا والاتحاد الدولي للاتصالات. فقد أصدت الإسكوا دراسات استقصائية حول وضع صناعة المحتوى الرقمي العربي في المنطقة العربية وأخرى تناولت نماذج الأعمال الجديدة لشركات المحتوى الرقمي العربي إضافة معالجة معمقة للموضوع في كافة إصدارات تقرير الملامح الإقليمية لمجتمع المعلومات. يرجى الاطلاع على الملحق ٢ الذي يبين دراسات وتقارير الإسكوا في مجال المحتوى الرقمي العربي.

---

[٣٨] http://time.com/3054279/50-best-websites-2014/

[٣٩] لا تقدم مجلة التايم معايير اختيارها لأفضل المواقع ويبدو ذلك نوعاً من التقدير للمواقع التي برزت على نحو واضح في عام ٢٠١٤ وهو فيما يبدو تقدير تقوم به في نهاية كل عام. ومن معاينة المواقع المختارة يبدو أن هاجس من قام بالاختيار هو الفائدة التي تقدمها المواقع المختارة للناس في حياتها اليومية، وهذه المواقع في الحقيقة تحمل أفكاراً مجددة ولا تماثل مواقع موجودة سابقاً، وهو ما يتطابق مع مقدمة هذا الفصل.

كما أصدرت مؤسسة الفكر العربي التقرير التأسيسي للمحتوى الرقمي العربي (الواقع- الدلالات- التحدّيات)[٤٠] الذي يتناول معطيات ميدانية تتعلق بالمحتوى العربي بلغ قوامها أكثر من ٢٠ مليون وحدة، تمثّل ما أمكن الوصول إليه من محتوى مكتوب باللغة العربية على الإنترنت خلال فترة زمنية تمتدّ من العام ١٩٩٠ وحتى العام ٢٠١١، وتشمل المحتوى المكتوب باللغة العربية والمنتج من ٢٢ دولة عربية و٦٦ دولة مختلفة حول العالم. أما هدف التقرير فهو تقييم حالة المحتوى العربي وقد أشار التقرير إلى ضعف تواجد المحتوى العربي إضافة إلى بطء حركته ونموه. كما ويتحدث التقرير "عن مشكلة عميقة تتعلق المحتوى العربي وما يتعلق بها من معضلات تقنية بل وأخلاقية". ويخلص إلى مجموعة من التحديات والفرص كما يلي:

- *التحديات*: أولها الحجم القليل، وثانيها تحدي الجودة التي تضمن بقاءه، وتحدي التوزع الجغرافي المتمركز حالياً في دولتين عربيتين أو ثلاث إلى جانب المحتوى مجهول الجنسية. وكذلك تحدي الاختلال في الاهتمام المتركز في الترفيه والدين وتراجع قضايا التنمية، وتحدي التبعثر في المشروعات دونما إطار استراتيجي.

- *الفرص*: أما الفرص فتتلخص في أن العالم العربي يشكل سوق هائلة غنية بالتنوعات الاجتماعية والاقتصادية، وبالتالي فإن هذا العالم العربي قابل لاستيعاب المزيد من المحتوى بكل ما يمثل ذلك من فرص للعمل البيني. كما يمكن استغلال الاهتمام العالمي بالمحتوى العربي لتوليد قيمة مضافة. يضاف إلى ذلك أن المحتوى الرقمي يشكل فضاء جديد للإبداع والابتكار أمام شريحة واسعة من الشباب.

الإطار ٣- أسماء النطاقات العليا لرموز البلدان

من بين مؤشرات عديدة على ضعف اهتمام الدول بدورها في إدارة الإنترنت يمكن النظر في موضوع أسماء النطاقات العليا وخاصة تلك المتعلقة بالدول (السيادية)، أو ما يسميه البعض برموز البلدان على الإنترنت. فقد اتخذت دولٌ كثيرة موقفاً واضحاً في هذا الخصوص، مثل دول الاتحاد الأوروبي التي انتهجت منذ العام ٢٠٠٠ سياسة تشجع على أن يُشار إلى المواقع بلاحقة الدول التي تنشأ منها وفيها هذه المواقع لأن ذلك يعطي مصداقية أعلى للمواقع ومحتواها على عكس اللواحق العامة مثل (com.) أو (net.). وبالفعل نجد في دولة مثل فرنسا ازدياد عدد المواقع التي يتضمن اسم نطاقها الأعلى لاحقة (fr.) عن ثلاثة ملايين موقع، وكذلك فعلت دول أخرى مثل اليابان وتركيا (١,٣ مليون)[٤١] وغيرها. أما بالنسبة للدول العربية، فلا زالت أعداد التسجيل ضعيفة جداً بحيث بلغت ٦٩٥٣ في الجزائر، و ٤٥٥١ في الأردن، و ٥٧٠٠٠ في المغرب، و ٤١٠٠٠ في السعودية. إلا أن الإمارات العربية المتحدة تبذل جهوداً حثيثة لترويج أسماء نطاقاتها (ae. و.إمارات) حيث تخطى مجمل التسجيل فيها عتبة الـ ١٠٠٠٠٠، كما أن الإمارات ستكون المشغل لأسماء النطاقات العربية الجديدة (arab. و.عرب).

[٤٠] يمكن الاطلاع على معلومات حول الدراسة على الموقع:

http://www.arabthought.org/ar/التقرير-التأسيسي-للمحتوى-الرقمي-العربي-الواقع-الدلالات-التحدّيات#.Vheo5ivQDph

[٤١] http://www.webhosting.info/domains/country_stats/

ثالثاً- المحتوى ماضياً وحاضراً ومستقبلاً

ألف- أجيال الوب

كان المحتوى في بدايته، مطلع تسعينيات القرن الماضي، يمضي باتجاه واحد فقط، من الحاسوب إلى المتلقي، وهو ما عُرف بجيل الوب الأول W1.0 أو الوب الساكنة أو وب الوثائق، باستثناء مواقع البيع التي كانت تتيح لمستخدم الإنترنت القيام بما يلزم من عمليات لإتمام صفقة البيع. ومع مطلع الألفية الثالثة انتقل دور مستخدم الإنترنت من مجرد متلق إلى متفاعل مع المحتوى، عبر التعليقات والمشاركة. كما أصبح لمستخدم الإنترنت دور الناشر باستخدام أدوات ووسائط بسيطة مثل المدوّنات وصفحات التواصل الاجتماعي، وإمكانية تحميل الفيديو والصور عبر مواقع مختصة بذلك وهو ما عُرف بالجيل الثاني من الوب W2.0 ذات الصفحات الديناميكية. وقد أدى ظهور الجيل الثاني إلى تضاعف المعلومات وزاد من صعوبة البحث فيها للوصول إلى المعلومة المطلوبة.

وبالرغم من التطور الهائل الذي جرى في محركات البحث، إلا أنها تعجز في كثير من الأحيان أن تقدم للمستخدم ما يبحث عنه بالضبط من معلومات، أو أن توصله إلى الموقع الأكثر مطابقة لما يريد. فبعد كل طلب بحث عن معلومات ترد آلاف الاحتمالات مرتبة بحسب الكلمات المفتاحية المذكورة في الطلب وبحسب تردد الاستخدام للمواقع في حالة تشابه الكلمات المفتاحية. هذا بالإضافة إلى أن ما نبحث عنه قد يتوفر كلياً أو جزئياً في أكثر من موقع. ولكن لا شيء يضمن الوصول إلى المعلومة المطلوبة إذ قد تكون معايير الترتيب لمحرك البحث غير متلائمة مع معايير المستخدم. هذا مع الإشارة إلى عدم تمكن زواحف محركات البحث crawlers المختلفة من ملاحقة الجديد من المعلومات المتضخمة بشكل أسي والتي يعجز في نهاية الأمر أي محرك بحث عن الوصول إليها (انظر الإطار ٤).

وكما هو معروف فإن أدوات الوب التقليدية تقوم فقط بإخراج المحتوى وتقديمه للمتلقي مع الروابط المتوفرة (إن كانت لا تزال موجودة) إلى مواقع أخرى أو وصلات داخل الموقع نفسه دون أية معرفة بنوع المعلومات المقدمة وما تتضمنه الروابط من حيث المحتوى والدلالة. ومثل هذا يتحول إلى مشكلة عند البحث عن صور أو صور تتضمن معلومات غير مشمولة في الشرح المرافق للصورة على النسق الذي يقوم به غوغل، أو حتى النصوص المصورة كما هو الأمر في الكثير من الوثائق النصية المصورة التي لا يمكن إلا للمستخدم أن يقرأها.

ومن هنا نشأت الحاجة إلى جيل جديد من الوب هو الجيل الثالث W3.0، يسمى بالوب الدلالي semantic web [٤٢] التي تهدف إلى جسر الفجوة في التواصل بين التطبيقات الحاسوبية ومستخدمي الوب. ذلك أن تطبيقات الوب الحالية لا تربط البيانات أو المعلومات بالسياق الذي نشأت فيه، ومن ثم فإن هذه التطبيقات لا تميز بين المهم والأقل أهمية، ولا تحسن الربط بين البيانات أو المعلومات مختلفة المصادر لتستخرج منها المعلومات الأفضل.

[٤٢] American, Mai 2001 Tim Berneers-Lee, James Hendler, Ora Lassila, The Semantic Web, Scientific

الإطار ٤- محركات البحث المختلفة

في مواجهة عملية البحث العسيرة في مليارات الصفحات المنشورة على الوب ظهرت أفكار أخرى تدعو لأشياء مماثلة في صلبها لما يُبحث عنه في الوب الدلالي من حيث البحث عن تكامل المعلومات وفق نهج منظم في صياغة بيانات البيانات. وكذلك الربط بين بيانات البيانات هذه حيث يبدأ الانطلاق من قاعدة بيانات ما ومن ثم التحرك عبر عدد غير منته من قواعد البيانات المرتبطة فيما بينها دلالياً، أي التي تتناول الموضوع نفسه.

ولا يركز الوب الدلالي على الكلمات المفتاحية والتعابير وإنما على معاني المحتوى، ويتم هذا التركيز باستخدام لغة ذات مفردات أو نهج يسمح بتمثيل مفهوم عملي عبر ربط البيانات ببعضها بأشياء العالم الحقيقية بحيث يمكن للآلة أن تفهم المعاني الدلالي وأن تميّز بين "الأرز" عند الحديث عن الطعام و"الأرز" عند الحديث عن الشجرة.

فبدلاً من التعرف إلى صفحة مفيدة كما في محركات البحث المشهورة حالياً، سيكون التحرك إلى استخراج المعلومات، التي يبحث عنها المستخدم، من الصفحات كما في حالة محرك البحث http://www.cluuz.com وتقديمها للمستخدم. كذلك تحاول بعض محركات البحث الدلالية مساعدة مستخدم الإنترنت على تحديد عمليات بحث إضافية يمكن أن تكون أكثر فائدة لهم في سعيهم لبلوغ أهدافهم كما في حالة محرك البحث http://www.wolframalpha.com.

كما يمكن لمحرك البحث الجديد (الدلالي) أن يقدم أكثر من مجرد مقترحات، بل يمكن استخدام أجوبته لتلبية حاجة مستخدمه، كما في حالة محرك البحث http://trialx.com.

ومن أهم خصائص الجيل الثالث للوب نجد أن البحث عن المعلومات المفصلة وفقاً للطلب tailor-made وليس وفقاً لمعايير ترتيب محرك البحث كما هو الأمر حالياً، وفي ذلك أن البحث سيوافق الطلب الشخصي. كذلك، يكون البحث سياقياً وفق قواعد التفكير الاستنتاجي deductive. وعليه سيكون البحث الدلالي بالمعاني الآتية[٤٣]:

- التعامل مع التنوعات الصرفية الخاصة بالجمل وصيغ المفرد والجمع والسوابق واللواحق؛
- التعامل مع المترادفات بمعانيها الدقيقة وترابطاتها كما هو الحال مثلاً في كلمات مثل "استطبّ" و"تعالج" و"استشفى"؛
- التعامل مع التعميمات كما في حالة "مرض" التي هي تعميم للحمى والالتهاب وآلام الرأس ونقص المناعة؛
- التعامل مع تقابل المفهوم كما هو الحال عند الحديث عن "بما يعالج آلام الرأس" الذي يحتاج إلى التقابل مع مفهوم أن مرض الشقيقة ينتمي إلى مفهوم ألم الرأس بالمعنى الطبي؛
- التعامل مع استفهامات وأسئلة اللغة الطبيعية مثل: ماذا وأين وكيف ولما؛
- القدرة على التوجه في فقرة مستمرة إلى الجملة الأهم فيها أو خلاصتها؛
- القدرة على طرح الأسئلة بحرية دون الحاجة لمصاغات format خاصة مثل الأقواس المزدوجة أو المؤثرات البولانية Boolean Operator من نوع "أو or" و "و and"؛

[٤٣] https://flatworldbusiness.wordpress.com/flat-education/previously/web-1-0-vs-web-2-0-vs-web-3-0-a-bird-eye-on-the-definition/

- القدرة على العمل بالاستناد إلى الإحصائيات أو سلوك المستخدم أو أية وسائل اصطناعية أخرى كما هو الحال في محركات البحث الحالية؛
- قدرة هذه المحركات على استشعار أدائها الخاص.

هذا الجيل هو مشروع قيد التنفيذ، حدث فيه تقدم، والعمل فيه أمر مستمر، وتحققه سيحدث تغيراً نوعياً في الوصول إلى المعلومة والتطبيقات المختلفة التي ستنجم عن ربط البيانات بواسطة بيانات البيانات metadata.

وخلال الحديث عن الجيل الثالث يجري الحديث عن الجيل الرابع W4.0 حيث ستتكيف الوب بما يسمح بربط كل أجهزة الاتصال النقالة بالإنترنت والاستفادة من محتوى الوب بالكامل. وهذا أمر تحقق تقريباً بالكامل، لذا جرى الاستعاضة عنه بالجيل الخامس من الوب W5.0. وهو جيل الوب المفتوحة والمترابطة والذكية أو الوب الانفعالية، وهذه الوب لا تزال مجرد فكرة مشروع يُطمح عبرها أن يتمكّن المستخدمون من التواصل فيها مع الآلة كما يتواصلون فيما بينهم.

وهذه الأجيال من الوب تطرح مشكلات عديدة أهمها مشكلة التشغيل البيني ومشكلة الدلالة اللغوية التي هي موضوع بحث علمي يجري على اللغات، من بينها العربية، وتتفاوت درجة التقدم في هذا البحث بين لغة وأخرى، كما تتفاوت الصعوبات الدلالية بين لغة وأخرى.

باء- الوب الدلالي

١- طبقات الوب الدلالي

الوب الدلالي هي توسيع للوب الحالي، تأخذ فيها المعلومات معان محددة، ممكّنة بذلك الأفراد من التفاعل على نحو أفضل مع الحاسوب. لذا على الحاسوب في الوب الدلالي أن يتعامل مع المعاني. وهذا يتطلب توفر المعالجة الحاسوبية للغة الطبيعية من حيث الاسترجاع التقليدي للمعلومات. كما أن محتوى الوب من اللغات الطبيعية يحتاج لأن يرافقه على نحو صريح معطيات دلالية لترميز المحتوى الذي يمكن للحاسوب قراءته وفهمه على نحو صحيح وذلك باستخدام أدوات برمجية تعطي وصفاً دلالياً لصفحة الوب بدلاً من الوصف بالكلمات المفتاحية كما هو الحال في جيل الوب الأول. أي أن الوب الدلالي يتضمن دلالية صريحة explicit web، حيث تكون المعلومات صريحة عن طريق تمثيل معياري للمعرفة المرتّبة منطقياً. وهذا ما يجري عن طريق ضمّ المعاني في شجرة تمثل مفهوماً، وهو ما يسمى ببناء الأنطولوجيا المرجعية. وهو الأمر الذي سيمكّن الحاسوب من معالجة معنى أو دلالة المعلومة آليّاً، وأن يربط ويكامل معلومات أو بيانات مختلفة الأنواع، وأن يشتق معلومات ضمنية من المعلومات الصريحة الموجودة بطريقة آليّة أيضاً.

يحدد ائتلاف الشبكة العنكبوتية العالمي خمس طبقات للوب الدلالي[٤٤]:

[٤٤] http://www.w3.org/standards/semanticweb/

- **الطبقة الأولى** (القاعدية) في الوب الدلالي هي ما هو منشور في الوب الحالية من بيانات وتواريخ وعناوين وغير ذلك، ولكن بنسخته الدلالية، حيث ستسمح أدوات الوب الدلالي (مثل RDF التي تمثل إطاراً لوصف المصادر، وهي تسمح بوصف البيانات في مصدر ما وربط هذا المصدر بمصادر أخرى) من ربط البيانات عن طريق بيانات البيانات metadata المرافقة للبيانات بما يمهد ويسمح للانتقال إلى الطبقة الثانية بسلاسة؛
- وفي **الطبقة الثانية** يجري تنظيم البيانات وبناء مخازن المفردات أو الأنطولوجيات[٤٥] وذلك بواسطة أنظمة تنسيق المعرفة البسيطة حيث يجري إغناء البيانات بدلالات ومعانٍ إضافية تتيح للناس وللحواسيب استفادة أكبر من البيانات وحيث يمكن للحواسيب فهم هذه البيانات وترابطاتها. وقد أصبحت أدوات بناء مخازن الأنطولوجيات هذه معيارية تماماً؛
- وفي **الطبقة الثالثة** تأتي مرحلة استخراج ما في البيانات من معلومات ومعارف عن طريق طرح الاستفهامات والأسئلة. ذلك أنه ينظر إلى الوب الدلالي على أنه قاعدة بيانات شاملة؛
- أما **الطبقة الرابعة**، فهي المرحلة الفنية الأخيرة التي يجري فيها الاستنباط والتفكير في البيانات بواسطة قواعد محددة؛
- و**الطبقة الخامسة** هي طبقة التطبيقات مثل تطبيقات العناية الصحية وتطبيقات علوم الحياة، والحكومة الإلكترونية والطاقة.

٢- انتقادات الوب الدلالي

من أهم الصعوبات[٤٦] التي يمكن أن تواجه الوب الدلالي هي مصداقية المعلومات الخاصة ببيانات البيانات في ظل تواجد طرق كثيرة لوصف الشيء نفسه. إضافة إلى أن البيانات نفسها قد تصبح خاطئة بمرور الوقت. وانتقاد آخر يوجه للوب الدلالي يتمثل في سهولة مراقبة الحكومات للمحتوى، أي أن الوب الدلالي ستكون عاملاً آخر ضد الخصوصية.

جيم- اللغة العربية والوب الدلالي

ترتبط الطبقات الأربعة الأولى المذكورة سابقاً في مكوناتها وأدواتها بما يسمى **"معالجة اللغات الطبيعية"**، وهو أمر في نهايته مرتبط بكل لغة وخصوصياتها من حيث المعالجة وإن كان هناك اتفاق على المبادئ والمفاهيم. ولا يبدو أن محتوى اللغة العربية أحرز تقدماً يجعله متأهلاً للدخول في الوب الدلالي. كما أن الأدوات المتوفرة حالياً، في اللغات الأخرى كالإنكليزية، لا تدعم فعلياً اللغة العربية بما يسمح بتحويل المحتوى العربي الحالي إلى محتوى دلالي[٤٧]. ويذهب آخرون للقول بأن اللبنات الأولى (النصوص المنشورة على الوب الحالية) تفتقر إلى المعيارية من حيث استخدامها

---

[٤٥] تعرف القواميس اللغوية الأنطولوجيا على أنها "علم الكينونة"، بمعنى دراسة الخواص العامة لما هو موجود. يترجم ذلك في المعلوميات بأن الأنطولوجيا هي المجموعة المرتبة للمصطلحات والمفاهيم الممثلة لمعنى حقل من حقول المعلومات، والعلاقة بين هذه المفاهيم. تسمح الأنطولوجيا بالمحاكمة الذهنية بخصوص مفاهيم مجال ما.

[٤٦] http://fr.wikipedia.org/wiki/Web_s%C3%A9mantique#Observations_sceptiques

[٤٧] Majdi Beseiso, Abdulrahim Ahmad, Roslan Ismail. of Arabic language Support in Semantic web; A survey http://www.semantic-web-journal.net/sites/default/files/swj110.pdf

لترميز متعدد[٤٨] للمحارف العربية مثل UTF-8 أو Windows-1256 أو ISO-8859-6 في الوقت الذي تستخدم معظم اللغات ترميز الـ Unicode الأكثر ملائمة لأدوات الوب الدلالي.

من جهة أخرى فإن المكتبة الأنطولوجية العربية لا تزال ضعيفة جداً، وما هو موجود حالياً في هذه المكتبة مكتوب بالإنكليزية (حتى نسبة ٤٩%) ولا تظهر العربية بأية نسبة محسوسة[٤٩]. انظر الإطار ٥.

ويكثر الحديث اليوم عن تطبيقات اختبرت وأثبتت فاعليتها[٥٠] وأخرى لن يكون وضعها في الخدمة بعيداً، في مجالات الصحة والأعمال والتعليم والحكومة الإلكترونية. كما أن كثيراً من المواقع المعروفة مثل الموسوعة الحرة أو مواقع التواصل الاجتماعي مثل الفيسبوك لا تستخدم في أدواتها إلا أدوات الطبقة الأولى من الطبقات الخمس المذكورة سابقاً. وبعض المواقع قطع شوطاً لا بأس به من الطبقة الثانية. في الوقت الذي لا تزال فيه مثل هذه التطبيقات في العربية بعيدة المنال، وهي حتى ليست في بداياتها. وهنا ستكون مشكلة المحتوى أعمق بكثير من تلك الموجودة حالياً والمتمثلة في نقص النصوص واستخدامات الإنترنت التي تشكل المحتوى. عندها سيكون الحديث عن فجوة في البناء الدلالي (معرفية) وليس عن مجرد فجوة رقمية أو فجوة محتوى.

الإطار ٥- بعض خصوصيات اللغة العربية

تنقسم الخصوصيات من حيث الشكل ومن حيث البناء والدلالة. فمن حيث الشكل تكتب اللغة العربية من اليمين إلى اليسار على عكس اللغات اللاتينية، كما أن الحروف منها ما تكتب متصلة مع ما بعدها وأخرى منفصلة. إضافة إلى خصوصية التشكيل، وهي خصوصية من حيث الشكل والدلالة. فقد يمكن للكلمة المفردة أن تفهم بعدة معان إذا رسمت بدون تشكيل كما هو الحال في فعل "قمع" الذي يمكن أن يُفهم المقصود منه على أنه فعل "قَمَعَ" أو أداة سكب السوائل في الآنية الضيقة الفتحات.

نادراً ما يُشكّل كتّاب العربية كتاباتهم لصعوبة ذلك، من حيث الأدوات والمعرفة، والاعتماد على أن السياق كفيل بإزاحة الغموض. ولكن ذلك ليس بالأمر المباشر. فجملة مثل "عظم الرجل" قد يكون المقصود منها "عظمُ الرِجْلِ" أو "عَظُمَ الرجلُ" أو "عُظِّمَ الرجلُ" وغير ذلك، وعندها نحتاج إلى سياق أعم تظهره الفقرة أو الموضوع المتناول. وهذا ليس بالمتاح دائماً في حالة الجمل المستقلة، مثل: "ما أطيب الطعام" التي لا يمكن أن تُفهم إلا بالتشكيل، فقد يكون المقصود منها الاستفهام عن الأطيب من الطعام "ما أطيبُ الطعامِ" أو يقد يكون الإعجاب بطيب الطعام أثناء تناوله: "ما أطيبَ الطعامَ".

مسألة أخرى تكمن في أن الشيء الواحد قد يشار إليه بعدة مترادفات مما يضفي صعوبة في البحث الدلالي في الوب كما هو الحال في كلمات مثل الحب والعشق والوجد والهوى والغرام والهيام. وكذلك العلاج والاستطباء والاستشفاء والتداوي. وهذا ما يزيد من صعوبة بناء الأنطولوجيات.

كما أن في اللغة العربية معان مختلفة لكلمات لها الرسم نفسه، كما هو الحال مثلاً في كلمة "ضرب"، التي لها أكثر من عشرة معان مختلفة يمكن التمييز بين باستخدام التشكيل من مثل: ضَرَبَ، ضَرُبَ، ضَرِبَ، ضَرَّبَ، ضَرْبُ،

---

[٤٨] The Arabic language and the semantic web: Challenges and opportunities; Hend S. Al-Khalifa & Areej S. 2007 KACST & Al- Wabil. The 1st International Sysmposium on Computers and Arabic Language & Exhibition, SCS

[٤٩] المرجع السابق و: Arabic Semantic Web Applications – A Survey; Aya M. Al-Zoghbya , Ahmed Sharaf Eldin Ahmed and Taher T. Hamza.

[٥٠] http://www.w3.org/2001/sw/sweo/public/UseCases/BBC/

ضَرَبٌ، ضُرُبٌ، ضَرِب، ضَرْب. ولو أخذنا الرسم الأول " ضَرَبَ" لكان له وحده أيضاً معان كثيرة تختلف بحسب السياق من مثل[٥١]: ضَرَبَ القلبُ تحرَّك، نبض. ضرَبَ شيئًا: أصابه وصدمه. ضرَب بالأمر عُرْضَ الحائط: أهمله، أعرض عنه، احتقره. ضرَب عُصْفورين بحجرٍ واحد: حقّق هدفين بعمل واحد. الخ.

أمر آخر يزيد في تعقيد مسألة البحث الآلي هو القواعد اللغوية والنحو[٥٢]، فالتذكير والتأنيث والجمع مثلاً ليست بالأمور الآلية التي يمكن أن تُعهد إلى حاسوب ليشتقها وإنما عليه العودة إلى معجم لغوي.

أما الأسماء في العربية، أسماء الأشخاص والمؤسسات أو المنتجات، فتمثل في حد ذاتها مشكلة على غير ما يمكن أن يحدث في اللغات الأوروبية. ذلك أن الأسماء في العربية هي في معظمها كلمات من العربية لها معنى لغوي (سمير، باسل، منصور، زهرة، ياسمين، الخ) وهذا يضيف مشكلة للتميز بين الاسم كفاعل وبين الكلمة كمعنى[٥٣]. أما في اللغات الأوروبية فالأمر مختلف حيث الأسماء بلا معنى في الغالبية العظمى (جورج، جان، كلوديا، برناديت، إليزابيت...)، إضافة إلى أن الأسماء تبدأ كتابتها بالحرف الكبير، لأسماء الأشخاص والمؤسسات والمنتجات، وهذا ما يُسهل تميزها والوصول إلى دلالة الكلمة دونما حاجة للبحث في سياقات أعم إن توافرت!

يمكن المتابعة في الحديث عن فرادة اللغة العربية في الكثير من الصفات والملامح التي تعقد أمر المعالجة اللغوية آلياً. هذا كله طبعاً إذا تركنا مسألة العامية والفصحى جانباً. كل هذا يجعل من مسألة البحث الدلالي في اللغة العربية أمراً معقداً. وهو الأمر الذي يتطلب جهداً هائلاً من أبناء اللغة العربية في ميدان البحث العلمي في معالجة اللغة العربية حاسوبياً، وخاصة عندما يكون التدخل مطلوباً من الحاسوب.

---

[٥١] http://www.almaany.com/ar/dict/ar-ar/%D8%B6%D8%B1%D8%A8/

[٥٢] حيث يصعب أن نجد نصاً خالياً من الأخطاء القواعدية والنحوية مما يغير في المعاني والدلالات.

[٥٣] مثل: "توقف محمود عن الكلام"، فقد يكون المقصود شخصاً اسمه محمود وقد يكون القصد صفة للتوقف بأنه محمودٌ!

رابعاً ـ التوصيات

ألف- بالنسبة للدول العربية

١. التوجه نحو وضع استراتجيات لبناء اقتصاد المعرفة في الدول العربية ونقلها إلى عصر التكنولوجيا والمعرفة بحيث تضمن ما يقتضيه ذلك من عوامل بشرية واقتصادية؛

٢. إيلاء المحتوى العربي في الاستراتيجيات العربية أهمية كبرى، خاصة في ظل علاقة محتوى اليوم بالمحتوى الدلالي (المعرفي) للغد، والتدريب على إنشاء هذا المحتوى؛

٣. وضع سياسات واضحة إزاء شبكة الإنترنت العالمية وإدارتها لجهة خطوط النقل وبوابات العبور وحواسيب التخديم المخزنة للمواقع ومحتواها؛

٤. اعتماد المبادئ الأساسية الأربعة لائتلاف الشبكة العنكوبتية العالمي W3C المذكورة سابقاً والاستفادة من القواعد الإرشادية التي وضعها الائتلاف لتحقيق هذه المبادئ ومن ثم صياغة القواعد الإرشادية الخاصة بكل دولة ومجتمع وخاصة بما يتعلق بالتطبيقات التي تتيحها الإنترنت؛

٥. التشجيع على استخدام أسماء النطاقات العليا لرموز البلدان الخاصة بالدول العربية وكذلك استخدام أسماء نطاقات الإنترنت بالحرف العربي؛

٦. زيادة الاهتمام بمعايير وجودة التعليم العام، وتوجيه عناية خاصة نحو كتابة المقال وبنائه.

باء- الجامعات ومراكز البحوث والدراسات ومجامع اللغة العربية

١. دعم وتعزيز البحوث الخاصة بمعالجة اللغات الطبيعية وخاصة بما يتعلق بمتطلبات الإنترنت ومتطلبات الأجيال القادمة من الوب؛

٢. المساهمة في بناء الأنطولوجيات الخاصة بالوب الدلالي والمتعلقة باللغة العربية، وقيام الجامعات على نحو خاص بدعم بيئتها المحيطة فيما يخص جوانب الوب الدلالي؛

٣. المساهمة في مراقبة تطبيق قوانين حماية الملكية الفكرية مراقبة مباشرة في المقالات والتقارير والأطروحات، ونشر التوعية حول الملكية الفكرية. وكذلك مراعاة استخدام البرمجيات بنسخها الأصلية؛

٤. توجيه عناية خاصة لكتابة الطلاب لتقاريرهم الدراسية من حيث المضمون واللغة. ذلك أن هذا النوع من الكتابة لا يتعرض له الطالب في التعليم العربي قبل المرحلة الجامعية إلا في النادر؛

٥. تشجيع الأساتذة والباحثين والمفكرين على كتابة ونشر المقالات العلمية باللغة العربية وعلى إنشاء صفحاتهم الخاصة عن طريق المدونات وشبكات التواصل الاجتماعي؛

٦. إتاحة منتجات البحث والتطوير الخاصة بالجامعات على الإنترنت وذلك نظراً لأهمية محتوى هذه المنتجات معرفياً ولكون منتجاتها معيارية في مجالاتها.

جيم- المنظمات الدولية والإقليمية

١. بيان أهمية البحث والتطوير والابتكار في التنمية، ومساعدة الدول الأعضاء في وضع سياسات واستراتجيات للانتقال نحو مجتمعات المعرفة؛

٢. مساعدة الدول الأعضاء في صياغة سياساتها الخاصة في إدارة البنية الأساسية الدولية للإنترنت من الشبكة الحاملة وبوابات عبور ومخدمات بما يضمن مصالح هذه الدول والمحافظة عليها؛

٣. دعم الدول الأعضاء في تطبيق الملكية الفكرية ونشر استخدامها في المجتمع؛

٤. رفع الوعي حول التغيرات الاستراتيجية في بنية المحتوى وتأثيراتها على هذه الدول، كما هو الأمر حالياً بالنسبة للوب الدلالي.

دال- المجتمع المدني

١. المساهمة في نشر الوعي الخاص بقوانين الملكية الفكرية؛

٢. المساهمة في إغناء المحتوى العربي عن طريق عرض مختلف أنشطته عبر مواقع خاصة بهذا المجتمع المحلي بحيث تبنى وفق المعايير الواردة في هذه الدراسة؛

٣. زيادة الاهتمام بمعايير وجودة التعليم في المؤسسات التعليمية التابعة للمجتمع المدني، وتوجيه عناية خاصة نحو كتابة المقال وبنائه؛

٤. المساهمة في تشجيع الكتاب والمفكرين والشخصيات العامة على التواجد على الوب عبر مدونات خاصة بهم ينشرون عليها إنتاجهم وآراءهم، لما يمثل ذلك من محتوى رصين وغني للإنترنت.

هاء- شركات تصميم وتنفيذ المواقع وصيانتها

١. الاهتمام بصناعة المحتوى الرقمي العربي لأهمية هذه الصناعة للمنطقة العربية ولبناء اقتصاد المعرفة؛

٢. العمل وفق معايير إنشاء المواقع المذكورة في هذه الدراسة؛

٣. عرض مختلف الخدمات التي يمكن أن تساعد زبائنها على الارتقاء بمواقعهم من حيث الجانب الفني والاستثماري.

الملحق ١- أبرز محاور البحث في معالجة النصوص والكلام والجهات المعنية

أ- معالجة الكلام

- *تركيب الكلام*: هناك أبحاث في العديد من الدول العربية لتطوير نظم تركيب للكلام، ومنها في المغرب (جامعة محمد الخامس في الرباط)، وفي الجمهورية العربية السورية (المعهد العالي للعلوم التطبيقية والتكنولوجيا: إصدارات متتالية آخرها يتضمن تنغيماً طبيعياً)، وفي الجزائر نظام "منطيق" للباحث طه زروقي [٥٤]، وفي السعودية مدينة الملك عبد العزيز للعلوم والتقنية، وفي مصر (شركات IBM وصخر و شركة التطوير البحثي الدولي RDI). ويوجد مجموعة من البرمجيات المتاحة التي تساعد في تركيب الكلام باللغة العربية ضمن مشروع MBROLA [٥٥]؛
- *تعرف الكلام*: في لبنان (جامعة البلمند)، وفي السعودية مدينة الملك عبد العزيز للعلوم والتقنية، وفي مصر (شركات IBM وصخر و RDI). ثمة اهتمام أيضاً في دول أخرى مثل فرنسا (LIMSI)، والولايات المتحدة (BBN Technologies, Microsoft). يوجد مجموعة من البرمجيات المتاحة التي تساعد في تعرف الكلام مثل SPHINX [٥٦]؛
- *تعرف المتكلم*: في لبنان (جامعة البلمند) والجمهورية العربية السورية (المعهد العالي للعلوم التطبيقية والتكنولوجيا). ويتيح المعهد الوطني للمعايير والتقانة National Institute of Standards and Technology (NIST) مدونات صوتية خاصة بتقييم نظم تعرف المتكلم تتضمن متحدثين باللغة العربية.

ب- معالجة النصوص

- *تعرّف الخط العربي*: تملك شركة صخر [٥٧] برنامجاً لتعرف الخط العربي، المطبوع أو بخط اليد، بدقة قدرت بـ 99.8% للوثائق عالية الجودة، و96% للوثائق المنخفضة الجودة. كذلك، عملت شركة التطوير البحثي الدولي RDI على إنشاء نظام CleverPage لتعرف الخط [٥٨] حيث تصل فيه نسبة الخطأ في تعرف الحروف إلى 2.58% وفي تعرف الكلمات إلى 10.32%. كما تجري في الأردن (في كلية الهندسة التكنولوجية في جامعة البلقاء التطبيقية في عمان، وفي جامعة الأردن) مجموعة من الأبحاث على تحسين خوارزميات تعرف النصوص المطبوعة. كما يتوفر أيضاً محرك Tesseract [٥٩] المفتوح المصدر لتعرف الخطوط المطبوعة (Optical Character Recognition)، والذي يعمل على عدة نظم تشغيل، ويعتبر من أفضل محركات OCR المفتوحة المصدر، حيث أصبح يعالج اللغة العربية ابتداءً من الإصدار ٣. كما تتواجد العديد من المواقع المجانية التي تتضمن خدمة تعرف الخط العربي وتمكّن من استخلاص النصوص العربية من الصور مثل: newocr و i2OCR و OCRconvert؛

---

[٥٤] http://sourceforge.net/projects/arabic-espeak/?source=typ_redirect

[٥٥] http://tcts.fpms.ac.be/synthesis/mbrola.html

[٥٦] http://cmusphinx.sourceforge.net/

[٥٧] http://www.sakhr.com

[٥٨] http://www.rdi-eg.com/technologies

[٥٩] https://code.google.com/p/tesseract-ocr/

- *التحليل الصرفي*: ومن المحللات الصرفية محلل أرامورف Aramorph لشركة قاموس، ومحلل AraComLex للباحث محمد عطيه من مصر[٦٠]، ومحلل "الخليل الصرفي" الذي أعدته جامعة محمد الأول في وجدة – المغرب، والمحلل الصرفي للمعهد العالي للعلوم التطبيقية والتكنولوجيا، ومحلل SALMA من جامعة الأردن[٦١] وقد اهتمت بهذا المحور أيضاً دول أخرى مثل: فرنسا (جامعة ليون) والجمهورية التشيكية (جامعة تشارلز في براغ) بمحلل Elixir-FM والولايات المتحدة الأمريكية (جامعة بنسلفانيا وجامعة ستانفورد Stanford Morphological Parser)؛

- *نظام تحديد أقسام الكلام*: Part of Speech Tagging، مثل ArabicPOSLib من معهد قطر لبحوث الحوسبة QNRF والعديد من النظم في الدول الغربية، مثل MADA, AMIRA, ASMA, MADAMIRA من جامعة كولومبيا الأمريكية وStanford Arabic POS tagger من جامعة ستانفورد، وSAPA من مخبر LIMSI في فرنسا؛

- *تشكيل النصوص*: جرى العمل على تشكل النصوص في الأردن (جامعة عمان)، والجمهورية العربية السورية (المعهد العالي للعلوم التطبيقية والتكنولوجيا)، وفي لبنان (جامعة البلمند)، وفي الجزائر نظام "مشكال"[٦٢] المفتوح المصدر للباحث طه زروقي، وفي المغرب (جامعة محمد الأول) نظام الخليل للتشكيل؛

- *التحليل النحوي*: تجري العديد من الأبحاث بهدف تطوير محللات نحوية مثل Arabic TurboParser من Carnegie Mellon University في قطر، وفي الولايات المتحدة الأمريكية (جامعة ستانفورد) Stanford parser؛

- *نظام تعرف المسميات*: (أي أسماء العلم، الأماكن، المنظمات، وغيرها) Named Entity Recognition، مثل (AQMAR من Carnegie Mellon University في قطر)؛

- *نظام تحديد معاني الكلمات*: Word Sense Disambiguation ثمة أعمال في مخبر Laboratoire[٦٣] LaTTiCe؛

- *نظام تصحيح الأخطاء الإملائية*: مثل نظام آيسبل المفتوح المصدر للباحث طه زروقي من الجزائر[٦٤]؛

- *الترجمة الآلية*: هناك العديد من البحوث الجارية بهدف تطوير نظم للترجمة الآلية، في الأردن (جامعة عمان)، وفي قطر (جامعة Carnegie Mellon ومعهد بحوث الحوسبة)، وفي مصر (شركات IBM وصخر و RDI). من جهة أخرى، يعمل معهد الدراسات المتقدمة[٦٥] في

---

[٦٠] http://www.attiaspace.com/

[٦١] https://www.academia.edu/9390823/SALMA_Standard_Arabic_Language_Morphological_Analysis

[٦٢] http://www.tahadz.com/mishkal

[٦٣] http://www.lattice.cnrs.fr

[٦٤] http://ayaspell.sourceforge.net/

[٦٥] http://www.undlfoundation.org

جامعة الأمم المتحدة في جنيف على مشروع لتطوير لغة تشبيك عالمية منذ عام ٢٠٠١ وهي لغة صنعية تسمح بالتعبير عن المعارف في اللغة الطبيعية ومن خواصها أنها تسمح بالترجمة في الزمن الحقيقي من لغة إلى أخرى، بحيث يجري تحليل النص في لغة وترميزه باللغة UNDL Universal Networking Digital Language[٦٦] ثم يجري تركيب هذه الرموز بلغة هدف أخرى، وتعتبر الجمعية العلمية الملكية في الأردن هي المسؤولة عن اللغة العربية فيه. كما تجري في جامعة كولومبيا الأمريكية أبحاث متعددة لإنشاء نظام للترجمة الآلية من اللغة العربية بلهجاتها المختلفة إلى الإنكليزية وبالعكس.

[٦٦] http://www.undlfoundation.org

ملحق ٢- تقارير ودراسات الإسكوا في المحتوى الرقمي العربي

رقم	التاريخ	العنوان	العنوان باللغة الإنكليزية	الرمز المرجعي	لغة الإصدار
١	١٩ حزيران ٢٠٠٣	اجتماع خبراء حول تعزيزالمحتوى الرقمي العربي	Expert Group Meeting on DAC Consolidation	E/ESCWA/ICTD/2003/WG/ 2/21	A
٢	٢٣ نيسان ٢٠٠٣	تعزيز وتحسين المحتوى العربي في الشبكات الرقمية	Enhancing and improving Arabic content on digital networks	E/ESCWA/ICTD/2003/3	A
٣	٣ أكتوبر ٢٠٠٣	مبادرة المحتوى العربية	Arab Content Initiative	E/ESCWA/ICTD/2003/10	A
٤	٢٩ آب ٢٠٠٥	المحتوى الرقمي العربي: الفرص والأولويات والتوجهات	Digital Arabic Content: Opportunities, priorities and directions	E/ESCWA/ICTD/2005/4	A
٥	٣١ أكتوبر ٢٠٠٥	نشرة تكنولوجيا المعلومات والاتصالات (٤)	ICT Bulletin No. 4	E/ESCWA/ICTD/2005/7	A
٦	٢٩ حزيران ٢٠٠٧	ورشة عمل افتراضية حول تعزيز صناعة المحتوى الرقمي العربي	Virtual Workshop on Enhancing DAC Industries	E/ESCWA/ICTD/2007/7	A
٧	ديسمبر ٢٠٠٩	الترويج لصناعة المحتوى الرقمي العربي من خلال الحاضنات التكنولوجية	Promotion of the Digital Arabic Content Industry through Incubation	http://css.escwa.org.lb/ICTD/ 1107/1-2.pdf	E
٨	٢٢ آذار ٢٠١٠	مسح للمحتوى الرقمي العربي: برمجياته وتطبيقاته وتقييم احتياجاته	A Survey on Digital Arabic Content: software, applications, and needs assessment	E/ESCWA/ICTD/2010/Tech nical Paper.1	A
٩	٢٢ آذار ٢٠١٠	نماذج لبرامج الأعمال وخطط التسويق والشراكات لصناعة المحتوى الرقمي العربي	Models for Business Plan, Marketing and Multi-stakeholder Partnerships for the DAC industry	E/ESCWA/ICTD/2010/Tech nical Paper.2	A
١٠	٢٢ آذار ٢٠١٠	تطوير المحتوى الرقمي العربي من خلال الاحتضان	Development of Digital Arabic Content: Incubation Requirements And Training Needs	E/ESCWA/ICTD/2010/Tech nical Paper.3	A
١١	أيار ٢٠١٠	مشروع تعزيز صناعة المحتوى الرقمي العربي في الحاضنات التكنولوجية	Promotion of the digital Arabic content industry through incubation	10-0206-May 2010-1000	A,E
١٢	١٣ نيسان ٢٠١١	آليات لبناء محتوى عربي تفاعلي متعدد الوسائط للمجتمعات المحلية: البوابات الإلكترونية المجتمعية	Mechanisms for community-driven interactive Arabic multimedia content	E/ESCWA/ICTD/2011/Tech nical Paper.3	A
١٣	٥ نوفمبر ٢٠١٢	وضع صناعة المحتوى الرقمي العربي في المنطقة العربية	Status of the Digital Arabic Content Industry in the Region	E/ESCWA/ICTD/2012/Tech nical paper.4	E
١٤	٢٠١٢	منشور: صناعة المحتوى الرقمي العربي	Brochure: The Digital Arabic Content Industry	E/ESCWA/ICTD/2013/Broc hure.1	A,E

١٥	٢٠ حزيران ٢٠١٣	نماذج الأعمال للمحتوى الرقمي العربي	Business Models for Digital Arabic Content	E/ESCWA/ICTD/2013/Technical Paper.1	E
١٦	٣٠ ديسمبر ٢٠١٣	مذكرة سياساتية حول المحتوى الرقمي العربي	Policy Note on Digital Arabic Content	E/ESCWA/ICTD/2013/Technical Paper.5	A
١٧	تموز ٢٠١٣	المحتوى الرقمي العربي *(دراسة للاتحاد الدولي للاتصالات بالتعاون مع الإسكوا)*	Digital Arabic Content Assessment Study *(ITU Study in collaboration with ESCWA)*	ISBN: 978-92-61-14946-8	A,E
١٨	٢٠١٥	كتيب: المحتوى الرقمي العربي تحفيز الابتكار وريادة الأعمال في المنطقة العربية	Leaflet: Digital Arabic Content Promoting Innovation and Entrepreneurship in the Arab Region	E/ESCWA/TDD/2015/Pamphlet.1	A,E
١٩	٢٠٠٣ - ٢٠١٥	ملامح مجتمع المعلومات في غربي آسيا ٢٠٠٣ إلى ٢٠١١ و ملامح مجتمع المعلومات في المنقطة العربية ٢٠١٣ إلى ٢٠١٥	Regional Profile of the Information Society in Western Asia 2003-2011 and Regional Profile of the Information Society in the Arab Region 2013-2015	http://www.escwa.un.org/wsis/profiles.html	A,E