مركز أوال للدراسات والتوثيق
AWAL CENTRE FOR STUDIES AND DOCUMENTATION

أرشيفو

ARCHIVO

العدد 5 - نيسان/ أبريل 2017

ثقافة أرشيفية

أرشيف الويب.. حفظ ذاكرة الإنترنت

فاطمة شاهين

من السّهل على أيّ شخص أن يأتي بخادم ويب، ويضعه في جهاز الكمبيوتر، ويتيحه للعالم. يمكن حدوث ذلك بشكل سهل جداً، ولكنّه هشّ جداً، لأنّ أيّ خطأ تقنيّ في التجهيزات قد يؤدّي إلى اختفاء هذا الموقع بما يحتويه، يؤكّد بروستر كاهلي، أحد المؤسّسين في منظّمة أرشيف الإنترنت1.

دخل العالم في القرن الحادي والعشرين عصر الإنترنت من بوّابته الواسعة، ودخلت معه ممارسات ثقافية جديدة واحتياجات متنوّعة وخدمات متناسبة، لتلبي متطلّبات هذا العصر وحاجات المستخدمين كافةً. وقد استطاعت شبكة الإنترنت بسعتها اللامحدودة وتعدّد مهامها واستخداماتها، أن تأخذ الحيّز الأكبر في حياة الإنسان، وأن تسهّلها بالجوانب العلمية منها والعملية والشخصية. لقد أصبحنا نوثّق حياتنا بشكل شبه يوميّ عبر وسائل التواصل الاجتماعي؛ الإنستغرام والفيسبوك وغيرها، وذلك من خلال نشر النصوص والصور ومقاطع الفيديو، وإبداء الرأي بإضافة التعليقات هنا وهناك، إلا أنّ عالم الشبكة الواسع والعميق واللامحدود، يخفي طبيعةً هشّةً تتطلّب بالتالي حلّاً يحفظ تاريخ الإنترنت، ويساعد على العودة إلى كل مكوّناته عند الحاجة. هذا الحلّ إنّما يتمثّل بعملية أرشفة الويب.

فما هي أرشفة الويب؟ وما هي أسبابها؟ كيف يتم هذا النوع من الأرشفة؟ وما هي أبرز الجهات التي تقوم بذلك في العالم؟

ما هي أرشفة الويب؟

تُعرّفُ مؤسَّسةُ الأرشيف الوطني البريطاني أرشفةَ الويب على أنها عمليّة جمع المواقع والمعلومات التي تتضمّنها شبكة الويب العالمية، وحفظها ضمن أرشيف.

أرشفة الويب هي عمليّة شبيهة بالأرشفة التقليدية للوثائق الورقية، حيث تتم عملية الاختيار والتخزين والحفظ والإتاحة. تتيح هذه العملية المواقع المؤرشفة، ليتم استخدامها من قبل الحكومات والشركات والمنظمات والباحثين والمؤرخين والجمهور العام. وكما هو الحال في الأرشيفات التقليدية، فإن أرشيفات الويب مجمّعة ومُعتنى بها من قبل مؤرشفين، يمكن تسميتهم في هذه الحالة «مؤرشفي الويب»2.

لماذا أرشفة الويب؟

أ - تغيّر الممارسات المعلوماتيّة

يومًا بعد يوم، وبوتيرةٍ متسارعة، تزداد مساحة الويب ضمن الممارسات المعلوماتية، ويظهر الإنترنت كأكثر قواعد المعلومات أهمّيّةً وسعةً وتلبيةً للاحتياجات. إنّ ما أنتجه الويب الثاني من تغيّر في ممارسات استخدام شبكة الإنترنت، أدّى إلى أن يصبح التفاعل سيّد الموقف، وبات المستخدم يضيف نصوصًا، ويكتب في صفحات الويب، ويبدي رأيه، ويتفاعل مع غيره من المستخدمين عبر الكثير من البرامج والتطبيقات، وبتنا نجد نصوصًا ووسائط متعدّدة في الشبكة قد لا تتوفّر خارجها. ومن هنا أهمية حفظها وأرشفتها لضمان الرجوع إليها.

ب - حفظ التراث الرقميّ

لقد تشكّل مع الإنترنت ما يمكن أن نسمّيه «التراث الرقمي». إنَّ نشر المزيد والمزيد من الوثائق بشكل رقمي، شكّل بحدّ ذاته تراثًا يتطلّب جمعه والحفاظ عليه، لضمان إجراء تعداد كامل للإنتاج الثقافي في البلاد. إن قوانين الإيداع في العديد من دول العالم تمّ تطويرها لتشمل، جمعًا وحفظًا، هذا النوع الجديد من الوثائق.إذًا، يأخذ التراث الرقمي الأهمية نفسها التي يأخذها التراث الورقي، وهنا تأتي أرشفة الويب كحاجة ملحّة تنهي مشكلة هشاشة الرقمنة، حيث تفرض طبيعة الوثائق الرقمية حلولًا جدّية لمنع اندثارها بمختلف أنواعها، من صفحات ويب ونصوص ووثائق سمعية وبصرية3.

كيف تتمّ عمليّة أرشفة الويب؟

باعتبار أنَّ الويب يتضمّن عددًا هائلًا من المواقع والمعلومات، يستخدم المؤرشفون الطرق الإلكترونية لجمع المواقع. تقوم عملية أرشفة الويب على جمع المواقع من أماكنها في الويب المباشر باستخدام برامج مخصّصة لذلك. هذا النوع من البرامج يُسمّى الزاحف crawler، حيث يقوم البرنامج «بالتجول» بين المواقع وعبر الويب، وينسخ المعلومات ويحفظها أينما ذهب. وتصبح المواقع المؤرشفة مع المعلومات التي تحتويها متاحةً على الخط المباشر (أونلاين)، كجزء من مجموعة أرشيف الويب. كما أنّ هذه الصّفحات والمعلومات التي تمّ جمعها، تصبح مرئيةً ومقروءةً وقابلةً للتصفّح، كما كانت عليه حين كانت مباشرةً في الويب، إلا أنها تكون هنا محفوظة على شكل لقطات ويب للمعلومات في مرحلة معينة من الزمن4.

من هي الجهات التي تقوم بأرشفة الويب؟

المكتبات الوطنيّة، الأرشيفات الوطنيّة، المنظّمات التي تُعنى بالمعلومات، المنظّمات التي

تُعنى بالتكنولوجيا، فضلًا عن غيرها من المنظّمات المختلفة والمتنوّعة. هذه الجهات معنيّة جميعها بالأرشفة المفصّلة لأهمّ ما يتضمّنه الويب من محتوى. إن البرامج والخدمات التجارية الخاصة بأرشفة الويب، أصبحت متاحة لكلّ المنظمات التي تحتاج إلى أرشفة محتوى الويب الخاص بها، وذلك لأهداف العمل، لحفظ التراث، لدواعٍ قانونية، وغير ذلك.

منظّمة أرشيف الإنترنت Internet archive

أ- نبذة عن المنظّمة

تُعتبر منظّمة «أرشيف الإنترنت»[5] من أهمّ المنظمات التي تُعنى بأرشفة الويب. يرد تعريفها في الصفحة الأولى لموقعها archive.org بأنها «مكتبة غير ربحية لملايين الكتب والأفلام والبرامج والموسيقى والمواقع الإلكترونية، وغيرها المزيد، بشكل مجّاني». تتيح المنظمة للباحثين والمؤرخين والأكاديميين والأشخاص ذوي الإعاقة والجمهور العام، الوصول الدائم إلى المجموعات التاريخية المتوفرة بصيغة رقميّة، وتهدف إلى بناء أكبر مكتبة رقمية في العالم.

تأسَّست المنظّمة في العام 1996، وتتّخذ من سان فرانسيسكو مقرًّا لها. تلقّت في بداياتها تبرعات بالبيانات من العديد من الجهات، أبرزها Alexa internet. في العام 1999، بدأت المنظّمة بالنموّ والاتّساع، ولا زالت كذلك، حتى وصلت اليوم إلى المحتوى الأضخم الّذي يتوفّر عبر موقعها archive.org في شبكة الإنترنت.

ب- محتويات موقع أرشيف الإنترنت

تضمّ منظمة أرشيف الإنترنت حاليًّا في موقعها 279 بليون صفحة ويب مؤرشفة، 11 مليونًا و331 ألفًا و882 كتابًا ونصًا، 3 ملايين و84 ألفًا و460 صورة متحركة (فيديو)، 3 ملايين و260 ألفًا و704 ملفات سمعية، مليونًا و323 ألف عرض تلفزيوني، 154 ألفًا و853 نتيجةً تتيح الوصول إلى ملايين البرامج وصور الأقراص المدمجة والتوثيق والوسائل المتعددة الوسائط، مليونًا و391 ألفًا و735 صورةً و168 ألفًا و837 ملفًا موسيقيًا.

يتيح الموقع إمكانية البحث في كل محتوياته، وإمكانية استعراض النتائج والوصول إلى الوثيقة المطلوبة، أيًّا كان نوعها. كما أنّ كل أنواع الوثائق المؤرشفة والمتاحة يمكن البحث عنها من خلال عنوانها، وتاريخ نشرها، واسم منشئها. كذلك، يوفّر إمكانية الوصول إلى 129 ألفًا و989 نصًا باللغة العربية من النصوص التي قام مستخدمو الموقع بتحميلها.

تشكّل الوثائق المؤرشفة للمكتبات المتنوعة حول العالم جزءًا مهمًّا من محتويات الأرشيف، فنجد

وثائق من المكتبات الأميركيّة، المكتبات الكندية، مكتبة روبارتس في جامعة تورنتو الكندية، مكتبة كاليفورنيا الرقمية، المكتبات الأوروبية وغيرها.

ج- Wayback machine

يقدّم موقع أرشيف الإنترنت مبادرة Wayback machine، وهي أداة تُستخدم للبحث عن صفحات الويب المؤرشفة. يتم استخدامها من خلال كتابة اسم الموقع الذي يتم البحث عن صفحة من صفحاته في خانة البحث، أو من خلال كتابة كلمة مفتاحية تعبّر عمّا يهدف الباحث الوصول إليه. وتقوم هذه الأداة باستدعاء كل النسخ المؤرشفة من المحتوى نفسه الخاص بالموقع، أو الكلمة المفتاحية التي كتبها الباحث بترتيبها الزمني6.

د- أرشيف الإنترنت في مكتبة الإسكندريّة

قامت منظّمة أرشيف الإنترنت في العام 2002 بتوقيع اتفاقية مع مكتبة الإسكندرية، تمتلك الأخيرة بموجبها نسخة احتياطية من الأرشيف، ويتم صنع موقعين متماثلين لأرشيف الإنترنت. وقد تم إهداء الجيل الأول من الأجهزة التي تعمل على أرشفة المواقع الإلكترونية لمكتبة الإسكندرية، كما تم إهداؤها أيضًا نسخة من أرشيف الإنترنت منذ العام 1996 وحتى العام 2001.

يُعتبر أرشيف الإنترنت اليوم أوّل مركز من نوعه خارج الولايات المتحدة الأميركية، ويتّسع حاليًّا لتخزين 4.9 بيتا بايت من البيانات المتنوّعة، وهو يُستخدم لأرشفة الإنترنت ومواد رقمية أخرى تتكوّن بأغلبها من المجموعة الهائلة من الكتب التي قامت المكتبة برقمنتها. ويمكن الوصول إلى أرشيف الإنترنت في مكتبة الاسكندرية عن طريق جهاز wayback machine، من خلال موقع archive.bibalex.org7.

تقول ماك أرثور، وهي موظّفة في منظّمة أرشيف الإنترنت، إنّ الرقمنة هي الحل اليوم لحفظ التراث المعرفي العالميّ، وتضيف أنَّ الحماية يجب أن تشمل أيضاً حفظ الصفحات الرقمية من الاختفاء والزوال8. لذلك، لا بدّ من العمل على تذليل العوائق أمام رقمنة المعرفة، ليس بالضرورة كبديل لما هي عليه حاليًا، بل كحلٍّ يحفظها من الاندثار، فاجتياز العوائق المالية والقانونية والتقنية سيؤدي حتمًا بالمكتبات وبكل الجهات المعنية بالمعرفة والمعلومات، إلى دخول العصر الرقميّ، والعمل على حماية التراث المعرفي بكل الطرق الحديثة التي أنتجها العقل البشريّ في عصرنا الحاليّ.

فاطمة شاهين: اختصاصيَّة في مجال المعلومات والمكتبات، ومفهرسة في مشروع النهوض بالمكتبة الوطنية اللبنانية. هي أيضًا مدربة في تنظيم المكتبات، وتعد حاليًا رسالة ماجستير في إدارة المعلومات.

للتواصل عبر الإيميل: shaheenfatima810@gmail.com