

المبادئ العامة في مجال استغلال ومشاركة المعلومات

عبد الكريم عبدون*

جامعة الجزائر – قسم المكتبات و علوم التوثيق

vac.abdelkrim.abdoun@univ-alger2.dz

تاريخ النشر: 2021 / 04 / 30

تاريخ القبول: 2021/02/27

تاريخ الإرسال: 2021/01/31

ملخص

تمثل علوم التوثيق و تطور مناهجها و أدواتها، الاجابة العلمية لإشكالية التدفق المعلوماتي و الوصول الى المعلومات. فقد أسست المبادئ الأولى للتوثيق في اواخر القرن التاسع عشر لهدف الوصول بأقل تكلفة الى الوثيقة المفيدة التي سيتم انتقاؤها من بين آلاف المنشورات المستجدة يوميا. و في منتصف القرن العشرين ظهر مصطلح الاعلام العلمي و التقني الذي يعنى بمحتوى الوثائق أي المعلومات العملية التي ستستخدم في البحث العلمي أو في التسيير المؤسساتي على حد سواء. و يمثل مفهوم الاعلام العلمي و التقني نقلة نوعية في علوم التوثيق كمجال علمي مفاده الوصول الانتقائي الى المعلومات الذي يعد المبدأ الذي يركز عليه تطور التوثيق و معالجة المعلومات و هو المبدأ الذي يفسر الاهتمام باليقظة المعلوماتية التي أصبحت في اواخر القرن العشرين مجالا قائما بذاته هدفه الانتقال من البحث عن المعلومات المفيدة الى المعلومات الضرورية و الاستراتيجية. و في كل هذه المراحل يمثل التطور التكنولوجي دعما و رهانا في الوقت نفسه لعلوم التوثيق التي تراهن في بداية القرن الواحد و العشرين على الاستحواذ على أحدث تكنولوجيا المعلومات كالواب الدلالي و البيانات الضخمة.

الكلمات المفتاحية: إعلام علمي و تقني، بيانات ضخمة، توثيق، تقييم، مشاريع وثائقية، واب دلالي، يقظة تكنولوجية

Abstract

Documentation science and the development of its methods and tools represent the scientific answer to the problem of information flow and access to information. The first principles of documentation were established in the late nineteenth century with the aim of at least costly access to a useful

document that would be selected from thousands of new documents which are published every day. In the middle of the twentieth century, the concept « Scientific and Technical Information appeared. Its object is to deals with the content of documents, that means, practical information that will be used in scientific research or in institutional management. The concept of scientific and technical information represents a qualitative leap in documentation sciences as a scientific field in the sense of selective access to information, which is the principle on which the development of documentation and information processing is based. This also explains the interest in information vigilance, which has become in the late twentieth century a field that aims to move from Searching for useful information into the necessary and strategic information. And in all these stages, technological development represents for documentation sciences, a support and a challenge at the same time. Now, this challenge concerns new technologies such as semantic web and big data.

Keywords: Big data, Documentation, Documentation projects, Information vigilance, Scientific and Technical Information, Semantic web

1. مقدمة

إن تطور المناهج و الطرق المتبعة لاستغلال و معالجة المعلومات مرتبط بعوامل تاريخية تتعلق بالتغير التكنولوجي و الاجتماعي . فمنذ اختراع الطباعة في القرن الخامس عشر إلى ظهور الواب الدلالي و الاهتمام بالبيانات الضخمة، توجب على محترفي علوم التوثيق، التفكير في مناهج و تقنيات مستحدثة في تفعيل البحث عن المعلومات و استغلالها بالطريقة الأمثل. سنحاول من خلال هذا العمل أن نحدد العلاقات بين التطور التكنولوجي و مناهج استغلال و معالجة المعلومات، مفترضين أن المسار التاريخي لعلوم التوثيق هو مسار نحو البحث الانتقائي عن المعلومات ؛ بمعنى أن تقنيات التوثيق و البحث عن المعلومات تبني على مبدأ الانتقاء و البحث عن الفعالية في الوصول إلى المحتوي. و كلما تزايد حجم المعلومات كانت الحاجة إلى الانتقاء اكبر مما يؤدي إلى تطور تقنيات استغلال و مشاركة المعلومات.

1. طرق الوصول إلى المعلومات

يجب القول أنه سواء تعلق الأمر بالبحث الآلي أو اليدوي ، فإن الهدف يكمن في البحث عن الدقة و الحد من الضجيج و الصمت (١). و لعل التطور التاريخي للعلوم وتوثيقها يعطينا نظرة واضحة عن صعوبة الوصول إلى المعلومات كلما زاد حجمها و انتشارها عبر المواقع. و كان الفلاسفة أو العلماء الموسوعيون أول من فكر في تقنين المعلومة العلمية عن طريق وضع تصنيفات للعلوم التي كانت في متناول الفرد نظرا لحجمها المتواضع و تواجدها في مواقع محدودة، حيث يمكن لمكتبة واحدة أن تسع جل العلوم المعروفة آنذاك.

بدأت الأمور تتغير مع اختراع الطباعة في القرن الخامس عشر حيث منحت دفعا جديدا لنشر وتوزيع المعلومات إلا أن هذا التطور ظل محدودا نظرا لتفشي الأمية و قصر التعلم على فئات معينة من المجتمع. و لم يعرف النشر تطوره الواسع إلا في أواخر القرن الثامن عشر بفضل تضافر عاملين متكاملين : الثورة الصناعية و ما جاءت به من تطورات في جل المجالات بما فيها مجال الطباعة، و كذلك انتشار التعلم بين طبقات المجتمع بعد أن كان قصرا على طبقات اجتماعية محدودة العدد. و بحلول القرن التاسع عشر، عرف العالم تطورا في الانتاج الفكري لم يُعرف له نظير في التاريخ. و لحاجة التحكم في هذا الحجم المتزايد من الوثائق، ظهرت بوادر علوم التوثيق التي تبحث عن أنجع الطرق للوصول إلى المعلومات. و قد كان لهذا المجال المعرفي الجديد عددا من الرواد من أمثال بول أوتلي (Paul Otlet) و هنري لافونتان (Henri La Fontaine) اللذان أنشأ سنة 1895 المعهد الدولي للبيوغرافيا الذي سعي بعد ذلك سنة 1930 بالمعهد الدولي للتوثيق.

و قد ساهم هذا المعهد في وضع الأسس العلمية للتحكم في المعلومات عن طريق حصر الانتاج الفكري، تصنيفه و تكشيفه. كما قام بتطوير التصنيف العشري للأمريكي ملفيل ديوي (Melvil Dewey) و نقله إلى تصنيف عشري دولي ظل سائدا لعشرات السنين.

و مع تطور التفكير الاقتصادي الذي تبع الحرب العالمية الثانية أصبح لزوما المرور من معالجة الوثائق إلى معالجة محتواها أي المعلومات. و في هذا السياق اعتمد محترفو التوثيق مجالا جديدا للاهتمام و الدراسة يتمثل في الاعلام العلمي و التقني حيث لقد استعمل هذا المفهوم لأول مرة في بداية الخمسينات مع إنشاء المعهد السوفيتي للإعلام العلمي و التقني (VINITI).

يمثل الإعلام العلمي و التقني مرحلة حاسمة في الاتجاه الانتقائي للوصول إلى المعلومات حيث إن الدقة و السرعة أصبحتا معياري النجاعة للتحكم في المعلومات.

و عرفت الثمانينات بروز مفهوم اليقظة التكنولوجية الذي يمثل بعدا آخر في الوصول الانتقائي إلى المعلومات حيث يتم الانتقال من الاهتمام بالمعرفة ككل إلى أوعية المعرفة أي الوثائق ثم إلى محتوى الوثائق أي المعلومات التي ستمر بدورها من معلومات مفيدة مع مفهوم الاعلام العلمي و التقني إلى معلومات ضرورية و استراتيجية مع مفهوم اليقظة التكنولوجية.

و قد تزامن هذا التطور مع بروز الانترنت و تطوير أدوات البحث التي أصبحت منذ سنوات 2000 تعتمد على المعطيات بدلا عن الوثائق مما يمثل نقلة جديدة في المسار الانتقائي للمعلومات. سنتطرق إلى موضوع الانترنت بعد أن نعرض المبادئ العامة للوصول إلى المعلومات.

2. المبادئ العامة للوصول إلى المعلومات

يعتبر الوصول إلى المعلومات المتراكمة أمرا مستعصيا في حال عدم وجود نظام متماسك لمعالجتها و استرجاعها عند الحاجة. و يقوم هذا النظام على مستويين متكاملين :

- معالجة المعلومات في كل منظومة على حدة.
- الربط بين المنظومات المعلوماتية و الاتصال فيما بينها.

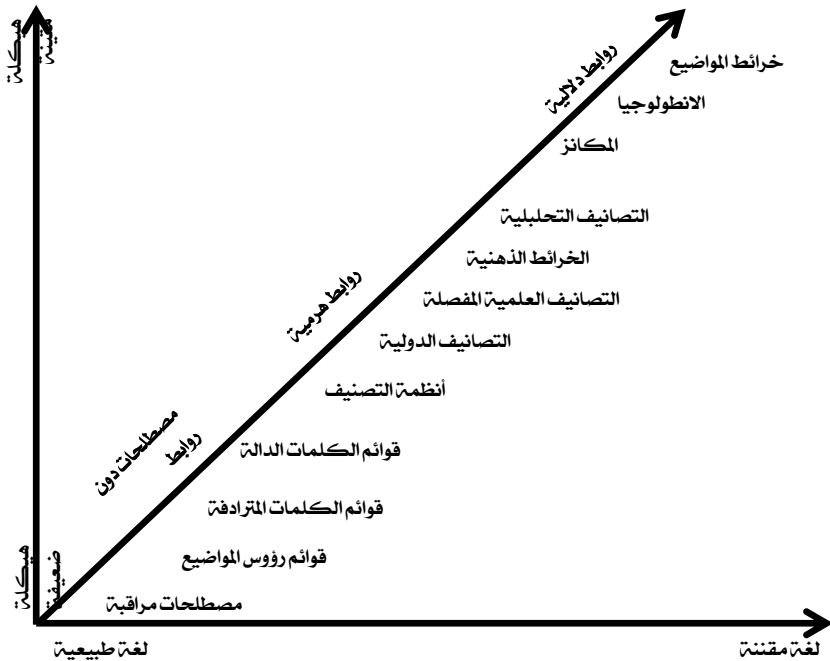
1.2. معالجة المعلومات

أول شرط أو مبدأ للوصول إلى المعلومة هو إعطاء هوية للوعاء الذي يحتويها، و تمثل بطاقة الفهرس أهم شكل من أشكال هوية الوثائق في حين يمثل محدد موقع الموارد (URL) الوسيلة الأمثل لتحديد هوية الوثائق المتواجدة عبر الانترنت في الواب "التقليدي" و يمثل معرف الموارد الموحد (URI) الوسيلة المتفق عليها في الواب الدلالي.

و يبقى تحديد مواقع الوثائق شرطاً أساسياً إلا أنه غير كافٍ للوصول إلى محتوى الوثيقة أي إلى المعلومات. و لهذا يلجأ محترفو الاعلام إلى وصف و تحليل المحتوى عن طريق عدد من التقنيات الوثائقية التي من بينها التصنيف بشقي أنواعها و التكشيف ، المراقب منه أو الحر. و يكمن التحدي هنا في تمتين التفاعل بين اللغة الوثائقية و اللغة الطبيعية لتأدية غرضين يمثلان الاشكالية الكبرى للوصول إلى المعلومات :

- تسهيل عملية البحث بالنسبة للمستخدمين.
- الزيادة من فعالية البحث إلى أبعد الحدود.

و لأجل تأدية هذين الغرضين تذهب الأبحاث في مجال الاعلام الآلي إلى جعل الآلة تؤدي وظائف البحث على أعلى مستوى حيث يمكنها فهم اللغة الطبيعية واستغلال أسئلة المستخدمين في صيغتها الطبيعية من أجل القيام بعملية البحث و استخراج نتائج قد تفوق تطلعات المستعمل نفسه. و يعتبر الواب الدلالي أقرب تكنولوجيا إلى بلوغ هذا الهدف.



شكل 1. التطور المنطقي والدلالي لتنظيم المعرفة (مترجم بتصرف عن :
)Pepper, Steve.- Closing SWOT Analysis. 2006. In : <http://www.ontopia.net>

2.2. الربط بين المنظومات المعلوماتية

ظهرت الحاجة إلى الاستغلال المتبادل للموارد المعلوماتية في نهاية القرن الثامن عشر كنتيجة للتدفق المعلوماتي، حيث لم يعد في وسع المنظومة الوثائقية أن تتحكم بمفردها في الكم الهائل من الوثائق وتلبية طلبات زبائنهم.

وقد ظهرت في هذا السياق عدد من المشاريع الكبرى لمشاركة الموارد (Resources Sharing) كالشبكات الوثائقية و الفهارس المشتركة. و مع تطور الانترنت أصبحت مشاريع تنظيم المعرفة تتمحور حول مشاركة البيانات (Data Sharing) خاصة منذ التسعينات مع بروز مفهوم البيانات الفوقية (Metadata) التي تمثل إحدى ركائز الواب الدلالي.

و تستلزم مشاركة البيانات شرطين أساسيين يتمثلان في التقييس و الاتفاقات بين الدول و/أو المؤسسات.

1.2.2 التقييس

يهدف التقييس إلى التغلب على الصعوبات التي تعيق التبادل بين الأنظمة المعلوماتية، و من بينها، يمكن ذكر العوائق التالية :

- تعدد تطبيقات الميئات
- التعدد اللغوي
- اختلاف مستويات و أنظمة التكشيف و الفهرسة الموضوعية

ظهر التقييس في معالجة المعلومات منذ أن جلت أهمية مشاركة الموارد الوثائقية في نهاية القرن الثامن عشر. و قد نشرت و هي مستعملة إلى يومنا هذا، عدة مقاييس دولية لتكشيف و فهرسة الوثائق، إلا أن تعميم التليماتية و الانترنت استوجب استحداث مقاييس جديدة لتمكين التبادل بين الحواسيب اعتبارا للعوائق المشار إليها. و يسرد في الجدول التالي بعض من أهم هذه المقاييس.

العنوان	المرجع
المكانز و التفاعل بين اللغات : Thesauri and Interoperability	ISO 25964-1
التفاعل بين اللغات Interoperability Between	BS 8723-4
التفاعل بين التطبيقات Interoperability Between	BS 8723-5
تصميم موحد لتبادل المعلومات Format pour l'échange	ISO 2709

جدول 1. بعض المقاييس في مجال تبادل المعلومات

2.2.2 الاتفاقات

لا يمكن للعمل التقني أن يبلغ فعاليته إلا إذا تبنته جل الأطراف الفاعلة في مجال حصر و توزيع المعلومات. و لهذا استوجب الأمر إنشاء منظمات حرفية و علمية على المستوى الدولي، تعنى بعدد من المهام لبلوغ هدف التعاون و الاستغلال الأمثل للتراث المعرفي، و من بين هذه المهام :

- إحداث الاتفاقات و العمل بها من طرف الناشطين في مجال تسيير المعلومات و المعرفة
 - البحث على وضع سياسات و قوانين وطنية للتكفل بالمكتبات و بمؤسسات الذاكرة بصورة عامة
 - إحداث المقاييس و/أو البحث على العمل بها
 - دعم و تشجيع المشاريع المهتمة بتطوير مجتمع المعلومات.
- و يمثل الجدول التالي بعضا من أهم المشاريع و الاتفاقات المعروفة في هذا السياق :

عنوان المشروع	صاحب المشروع
النموذج الموحد للمصطلحات الطبية : Unified Medical	المكتبة الوطنية الأمريكية
إشتقاق و نمذجة اللغات التوثيقية Faceted Application of	مكتبة الكونغرس الأمريكية
التفاعل بين التطبيقات Interoperability Between	BS 8723-5
الملف الألفبائي الموحد : RAMEAU	المكتبة الوطنية الفرنسية
إتجاهات التعاون و لتكامل بين المكتبات العامة و الأرشفة و	الاتحاد الدولي لجمعيات
إعلان اليونسكو الخاص بالتنوع الثقافي : Universal	UNESCO
الوثائق الصادرة عن القمة العالمية للمعلومات	الاتحاد الدولي للاتصالات
ربط البيانات المفتوحة : Linking Open Data	إتحاد شبكة الواب (W3C)

جدول 2. بعض المشاريع في مجال تبادل المعلومات

3. تطور الانترنت وتأثيره على معالجة المعلومات

يميز المتابعون لتطور الواب بين أربع أو خمس نقلات نوعية تمثل كل منها شبه ثورة في مجال الانترنت. و قد كان الواب أو ما يعرف بالواب 1.0 في بدايته ثابتا (statique) بالنظر لإمكانات التفاعل التي يمنحها لمستخدميه. و من أهم مستحدثاته : بوابات المحتوى، و بعض التجارة الخطية. و يتحدث المتابعون لهذا التطور عن الواب 4.0 الذي أصبح فيه الآلة شبيهة بالعقل البشري في معالجتها و تفهمها للمعلومات. و السؤال الذي يطرحه الوثائقي هو مدى تأثير هذه التغيرات على الممارسات المهنية و كيفية استخدامها في تحسين الخدمات الوثائقية.

1.3 النقلات النوعية

يميز المتتبعون لتطور الواب بين أربع أو خمس نقلات نوعية تمثل كل منها شبه ثورة في مجال الانترنت.

1.1.3 الواب 1.0

يرتكز على الوثائق و هو ثابت (statique) بالنظر لإمكانات التفاعل التي يمنحها لمستخدميه. و من أهم مستحدثاته : بوابات المحتوى، و بعض التجارة الخطية.

2.1.3 الواب 2.0

استعملت كلمة - واب 2.0 - أول مرة سنة 2004 أثناء ندوة نظمت من أجل البحث في التطورات التي عرفها الواب منذ نشأته ، إلا أن أوصاف الواب 2.0 تعود إلى قبل ذلك بسنين. و يرتكز هذا الواب على المستخدمين حيث أضحى نشر و تبادل المعلومات سهلا و في متناول الغير أخصائيين. و من أهم مستحدثاته : التجارة الخطية و شبكات التواصل الاجتماعي التي تعد أهم خاصياته حتى أن الواب 2.0 يعرف بالواب الاجتماعي.

3.1.3 الواب²

يمثل مرحلة انتقالية إلى الواب 3.0 الذي سيتجه إلى استغلال المعطيات. و من أهم مستحدثاته : الاعتماد على البيانات الفوقية (metadata) .

4.1.3 الواب 3.0 والواب 4.0

يختلف المتتبعون لتطورات الانترنت حول مضمون الواب 3.0 و 4.0 و الأرجح أنهما يمثلان النظرة التطلعية لما يكون عليه الواب في المستقبل. أما ما يُتفق عليه فإن مستجدات الواب في العشرية الأخيرة تتمثل في العناصر التالية :

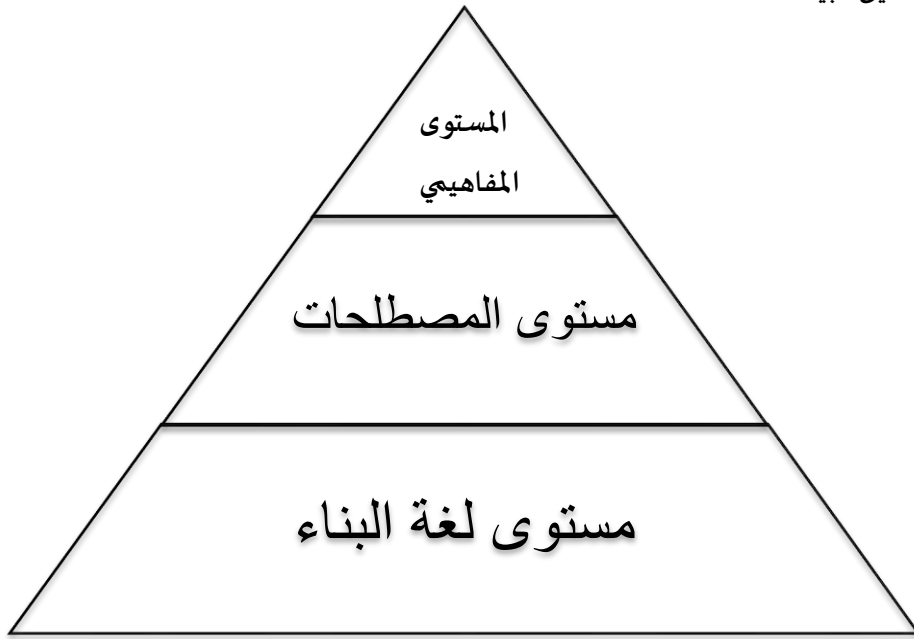
- زيادة التحكم في المعطيات و الربط فيما بينها
- تطور التجارة الخطية و التعامل عن طريق الواب في شتى المجالات
- الاعتماد على الذكاء الاصطناعي
- تطور تفهم الآلة للغة الطبيعية و بروز ما يعرف بالواب الدلالي
- الربط بين الآلات من غير الحواسيب و بروز ما يعرف بإنترنت الأشياء (internet des objets)

إذا كانت النسخ الوابية من 1 إلى 4 مروراً بالواب المربع (web squared) متعارف عليها و متداولة بين المستخدمين فإنها لا تمثل سوى اختصار لواقع تطور الواب الذي ينتقل من ربط الوثائق إلى روابط المعطيات. و قد خطط لهذه النزعة منذ أوائل التسعينات ، حيث أكد تيم بيرنرزلي خلال الندوة

الأولى للواب المنعقدة بجنيف سنة 1994 على ضرورة هذا الانتقال. و حررت في هذا الشأن سنتي 1997 و 1998 وثيقتان تعتبران كخريطة طريق لبلوغ هذا الهدف حيث تضمنتا المبادئ الأساسية للواب الدلالي. و لخصت هذه الرؤية في كتاب نشر سنة 1999⁽²⁾.

2.3 الواب الدلالي

يعرف الواب الدلالي على أنه " مشروع يستهدف جعل شبكة الواب أكثر قدرة على أن تفهمها أجهزة الكمبيوتر لتقديم مادة معالجة و ملائمة للمستخدمين ... و التكنولوجيا وراء ذلك هي لغة XML التي تسمح بإلحاق حقول الوصف بالمحتوى بالإضافة إلى إطار وصف المصدر (RDF) الذي يتيح قواعد البيانات الفوقية (Metadata) التي يمكنها استعمال حقول لغة الوصف (XML) في هيكلة ثلاثية لتمثيل البيانات.



شكل 2 . الأبعاد الثلاثة لتمثيل البيانات في الواب الدلالي

1.2.3 تقنيات الواب الدلالي

يرتكز الواب الدلالي على عدد من التقنيات أو التكنولوجيات أربعة منها أساسية :

- خراط المفاهيم أو الأنطولوجيا
- لغة الترميز القابلة للتوسع (XML)
- إطار وصف المصادر (RDF)
- لغة أنطولوجيا الواب (OWL)

2.2.3 خصائص الواب الدلالي

يحاول مصممو الواب الدلالي تمييزه بعدد من الخصائص، من ضمنها:

- الفرق بين الشكل و المحتوى.
- تقنين المعارف و نمذجتها.
- اعتماد الروابط الدلالية بين المعلومات.
- إمكانية تسيير المعرفة.
- الاعتماد على الانطولوجيا.
- العمل على التعرف على معاني الكلمات.

3.2.3 محركات البحث

تحاول محركات البحث في الواب الدلالي أن تأخذ بالاعتبار معاني و دلالات المصطلحات و ليس مطابقتها الشكلية فحسب كما هو الحال في محركات البحث التقليدية.

و هناك عدد من هذه الأدوات قد توقف بعضها و يزال بعضها في حيز التطوير، و يمثل الجدول 3 بعضا من أهم هذه المحركات.

المحرك	الموقع	سنة	صاحب المشروع	بلد
Swoogle	http://swoogle.umbc.edu	200	UMBC	و.م.أ
Mocassin	http://code.google.com/p/mocass	200	Google	و.م.أ
Factbites	http://www.factbites.com	200	Rapid	أست
Watson	http://kmi.open.ac.uk/technology	200	NeOn&OpenKn	بريطا
Biosemani	http://biosemantics.org	200	The	و.م.أ
Serelex	http://serelex.cental.be	200	UCL, BMSTU	روسي
Demo.pixo	https://demo.pixolution.de	200	Pixolution	ألمانيا
GeneView	http://bc3.informatik.hu-berlin.de	201	WBI	ألمانيا
Freebase	http://www.freebase.com	201	Google	و.م.أ
Tineye	http://www.tineye.com	201	Idée Inc	كندا

جدول 3. قائمة لبعض أهم محركات الواب الدلالي

4. البيانات الضخمة

تمثل البيانات الضخمة أبعد محلة في الوصول الانتقائي إلى المعرفة حيث تمثل المعطيات التي "يفوق حجمها قدرة الأدوات التقليدية في التخزين و المعالجة"³ و قد تكون منظمة (10%) أو غير منظمة (90%). و هي تتشكل من خلال النشاط الدائم المتواجد عبر الانترنات كالشبكات الاجتماعية و محركات البحث.

أما استغلالها فهو متعدد حيث تستخدم في كثير من المجالات كالطب و الشؤون العسكرية و في أنظمة البقطة الاستراتيجية كمدخل لدراسة التوجهات و تحديد حقول البقطة. و تستغل البيانات الضخمة عبر مرحلتين الخزن و المعالجة.

1.4 الخزن

أو ما يعرف بمستودع البيانات (Data Warehouse) وهو بنك للمعطيات وظيفته جمع و خزن و ترتيب البيانات المتوفرة عبر العديد من قواعد المعلومات لتحضير عملية الاستغلال.

2.4 المعالجة

أو ما يعرف بالتنقيب عن البيانات (Data Mining) ويقصد به كل الطرق و المناهج المستخدمة في تحليل المعطيات و استغلالها.

كخلاصة

يتضح من خلال ما ورد في هذه المساهمة المتواضعة أن الاتجاه هو منح الحواسيب نوعاً من "الذكاء" يمكنها من فهم المعلومات و معالجتها بشكل مقارب لما يقوم به العقل البشري. و ينتج عن هذا الاتجاه سؤال حول مستقبل علوم التوثيق و معالجة المعلومات.

بغض النظر عن ضرورة التكيف و الاستغلال الأمثل للتكنولوجيا المتوفرة في تحسين و تثمين الخدمات الوثائقية ، يكمن التحدي في هيكلة علم التوثيق ليضلل اللجوء الأمثل للباحثين عن المعلومات.

المراجع

(¹) يمثل الضجيج الوثائق أو المعلومات الواردة في نتائج البحث دون أن تفيد المعنى. أما الصمت فيمثل المعلومات المفيدة التي لم يتوصل إليها البحث.

(²) Berners-Lee, Tim.- Weaving the Web : the Ultimate Design and Ultimate Destiny of the World Wide Web.- New-York : Harper, 1999

(³) Mckinsey Global Institute.- Big Data : The Next Frontier for Innovation, Competition and productivity.- 2011