ترجمات

أدوار علم المعلومات في مجال علم البيانات الناشئ [1]

إعداد

جيرى مارشيونيني

ترجمة

أ. د. هاشم فرحات

أستاذ المكتبات والمعلومات
كلية الآداب – جامعة القاهرة

مقدمة المترجم

علم البيانات Data Science، هو علم وليد أو ناشئ ـ ظهرت معالمه وتجسدت في السنوات الخمس الأخيرة، ومن الملاحظ أن دائرة الاهتمام به تتسع كل يوم، حتى باتت كثير من الجامعات الغربية تقدمه بوصفه مساقًا تخصصيًّا فريدًا، سواء على مستوى المرحلة الجامعية الأولى، أم على مستوى مرحلة الدراسات العليا؛ وتحديدًا في مرحلة الماجستير، وهنا بالفعل برامج ماجستير متخصصة في هذا المجال، ويبدو أن ارتباطه بمعالجة ظاهرة البيانات الضخمة Big Data ـ تلك الظاهرة التي أراها وقد سيطرت على جل اهتمام الباحثين في مجالات عدة، وليس في مجال الحاسبات والمعلومات، هو الذي أكسب هذا المجال الوليد ذلك الزخم اللافت للنظر، وأعطى له هذا الاهتمام الكبير. غير أن الذي لفت انتباهي وأثار اهتمامي هو ذلك الجدل القائم حول طبيعة هذا العلم الوليد، ومجالات اهتماماته، وتبعيته التخصصية، وثم نقاش طويل أثير حول ارتباطه بعلم الحاسبات، أو بعلم المعلومات ـ أو أنها الجدلية الدائمة التي كثيرًا ما تثيرها خبايا كلمة " المعلومات " ـ تلك المظلة الكبرى التي جمعت وحوت الكثير من المتاعب لأهل الاختصاص ومن في حكمهم، وأثارت الكثير من المشكلات التي لا تزال رحى النقاش والاختلاف حولها مستمرة حتى اليوم. على أية حال لي عودة مفصلة حول قضية هذا العلم، وهويته، وتبعيته، ومجالات اهتمامه، غير أني أحببت أن أقدم للقارئ العربي هذه المقالة التي تعالج ـ أو ربما تثير ـ بعض الهموم حول ارتباط هذا العلم الوليد ـ علم البيانات ـ بمجال علم المعلومات. وأظن أنها ستجلي الصورة المشوشة حول بعض الحقائق التي سأترك للقارئ الكريم الحرية والفرصة لإعادة التفكير في بعض المفاهيم والرؤى التي كانت ولا تزال تثار حول حاضر علم

المعلومات ومستقبله، ويحدوني الأمل أن تترجم هذه الأفكار التي ستثيرها هذه المقالة إلى اتجاهات بحثية وجهود فكرية يتبناها شبابنا من الباحثين النشطين عن رؤية عصرية جديدة وواقعية لعلم المعلومات [المترجم].

المقالة الأصلية

ثمة نقاش دائر منذ أمد غير قصير حول الفروق بين علم المكتبات، وعلم المعلومات، والمعلوماتية Informatics، وحول أوجه الاختلاف بين هذه المجالات وعلم الحاسب، والمساحات التي تتداخل فيها مع هذا العلم. ويثير مصطلح علم البيانات Data Science الذي يظهر اليوم، تساؤلات واستفسارات مثيرة للتفكير حول علاقته بمجالات علمية أخرى، ومدى اختلافه عنها. ولأغراض هذا النقاش الذي نطرحه في هذه الصفحات، أرى أن علم المعلومات مصطلح عام ينضوي تحته علم المكتبات، والمعلوماتية، ويركز على أوجه الاختلاف والتشابه بين تلك التخصصات التي يتناول كل منها علم البيانات. فبشكل عام يمكن القول إن علم المعلومات يعنى بنشأة المعلومات وتدفقها واستخدامها وحفظها، فيما يتناول علم الحاسب الخوارزميات وتقنيات العمليات الحاسبية. أما علم البيانات بوصفه مفهومًا، فينشأ من تطبيقات الدراسات الحالية في مجال المقاييس، وأساليب تمثيل البيانات وتفسيرها، والإدارة، كل ذلك بغرض التعامل مع قضايا مجالات التجارة والصحة والبيئة والحكومة وغيرها من المجالات الأخرى. ومن المعروف أن كل مجال من مجالات الجهود البشرية يحرص وبقوة على الاستفادة من قدراتنا المتسارعة في التطور على الحصول والمشاركة والتحليل لذلكم السيل المتدفق من البيانات الناتجة عن الأنشطة الطبيعية والصناعية. وفي النهاية ـ يمكن القول ـ إن علم البيانات كسب أهميته من قدرته على مساعدة الأشخاص على اتخاذ قرارات أفضل، كل في مجال اهتمامه، أي كيف يطبق المتخصصون في الطب أو التجارة أو الحكومة أو البحوث مبادئ هذا العلم وتقنياته في معالجة المشكلات الخاصة بهم.

وفي كتابهما " النموذج الفكري الرابع: الاكتشافات العلمية ذات البيانات المكثفة The Fourth Paradigm: Data-Intensive Scientific Discovery "،(Hey, Tansley, Tolle, 2009)، تناول كل من هاي Hey و تولي Tolle، العديد من الأفكار التي سوقت لعلم البيانات. وقد اشتمل هذا الكتاب على عدة فصول تتناول أربعة مجالات تطبيقية مختلفة هي: البيئة، والصحة، والبنية التحتية العلمية، والاتصال العلمي، ويقدم أمثلة على كيفية استخدام مجموعات البيانات الضخمة Datasets للخروج باكتشافات في هذه المجالات المعرفية المختلفة. ومن الملاحظ أن عنوان الكتاب مبني على ادعاء جيم جراي Jim Gray بأن النماذج الفكرية العلمية الثلاثة الموجودة، وهي: النموذج الفكري التجريبي، والنموذج الفكري النظري، والنموذج الفكري الحاسوبي ـ قد تعززت بهذا النموذج الفكري الجديد القائم على استكشاف البيانات Data Exploration. ومن المعروف أن عملية استكشاف

البيانات تقوم بشكل أساسي على الوصول إلى سيل كبير من البيانات المتدفقة، ومعالجتها، وتحليلها. ولا يخفى أن عملية معالجة البيانات تدخل بشكل مباشر في نطاق علم المعلومات، ولدى مدارس المعلومات الكثير مما يمكنها أن تسهم به في مجال علم البيانات، مع ما سيشهده من تطور خلال السنوات القادمة.

وقد بات شائعًا الآن وصف علم البيانات وتمييزه بسلسلة من الكلمات التي تبدأ بحرف V، وهي : الحجم Volume، والسرعة Velocity، والتنوع Variety، والمصداقية Veracity، والقيمة Value، والتمثيل التصويري للبيانات Visualization (Zikopoulos & Eaton, 2011; McAfee & Brynjolfsson, 2012)؛ حيث أصبح يتعين علينا اليوم التعامل مع كميات هائلة من البيانات (مثل البيتابايت petabytes، والإكسابايتexabytes)، التي تتدفق بمعدلات عالية (جيجابت في الثانية gigabits/ second)، من مجموعة متنوعة من المصادر (مثل أجهزة الاستشعار والاستقبال sensors، ووسائل التواصل الاجتماعي، والمعاملات الإلكترونية، والكيانات البشرية أو الاصطناعية)، وبأشكال عديدة ومختلفة (مثل الأرقام والنصوص والأصوات والرسوم والفيديو).

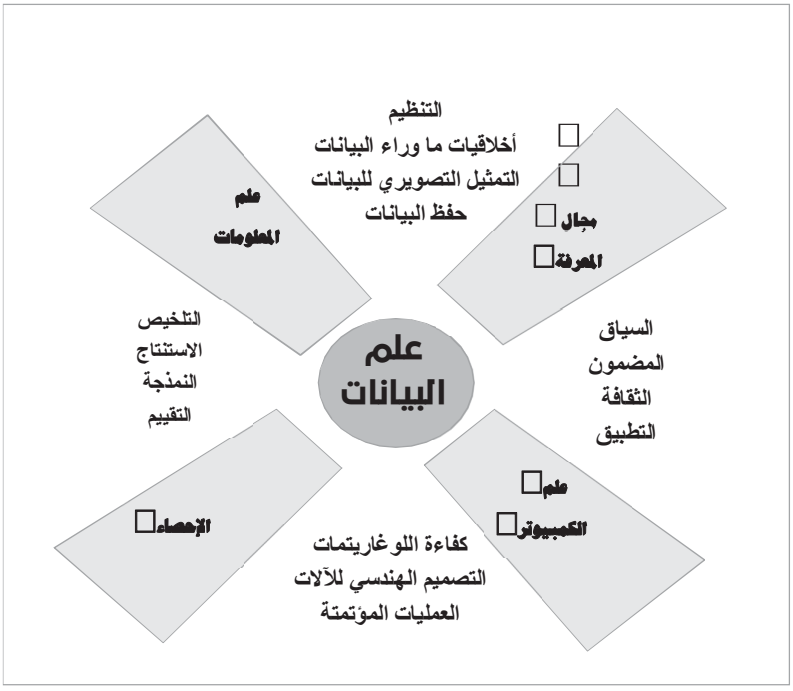
ومن الملاحظ أن الجهود المبكرة للاستجابة للأوصاف الثلاثة الأُول (الحجم والسرعة والتنوع) قد دفعت إلى تحفيز العمل في تقنيات التخزين الموزع أو الشبكي للبيانات (مثل شبكات البيانات Data Grids، والتخزين السحابي Clouds)، وعمليات التحسيب Computation (مثل المعالجة المتوازية Parallel Processing[3]، ونظام التخزين والمعالجة (MapReduce) [4]، هذا فضلًا عن بعض التقنيات الأخرى المماثلة عبر ما يسمى ببحيرات تخزين البيانات[5]Data Lakes)، يضاف إلى ذلك الأساليب الحاسوبية المختلطة لدمج البيانات المنظمة وغير المنظمة.

وهناك اعتراف متزايد بأن أساليبنا أو تقنياتنا لن تكون مجدية إلا إذا كنا نتعامل مع بيانات سليمة Good Data (وليس فقط مجرد كم كبير من البيانات)، وكذلك إذا كانت الأساليب التحليلية مناسبة للبيانات وللمشكلات المطلوب حلها. وفي هذا الصدد يمكن الرجوع ـ على سبيل المثال ـ إلى ما قدمه توفيكسي Tufekci (2014) في دراسته النقدية لتكدس Hype البيانات الضخمة، والتحديات المنهجية في التعامل معها، وكذلك إلى ما قدمه بورجمان Borgman (2015)، أو جونسون Johnson (2016) حول تعرف نماذج لكيفية تضمين البيانات في الأنشطة الاجتماعية.

وعليه يمكن القول ـ وبشكل قوي ـ إن علم المعلومات، يتناول علم البيانات، ذلك إذا أخذنا بعين الاعتبار الاهتمام القوي بالدورة الكاملة لحياة البيانات، وليس فقط الاهتمام بتخزينها وتحليلها، وهذا التركيز المتكامل تبرز أهميته بشكل خاص مع عنصري "المصداقية" و"القيمة"، المعروفتين في سياق علم البيانات.

ومن المعلوم أن علم البيانات في هذه المرحلة المبكرة من تطوره ـ وكما هي الحال مع أي علم ناشئ ـ يستمد نظريته وممارساته من القواعد المعرفية القائمة. ويوضح الشكل التالي، كيفية نشأة علم البيانات من صلب أربعة قطاعات رئيسية؛ ثلاثة مجالات منها تبدو متمايزة عن بعضها البعض ـ لكن ما بينها من ارتباطات لا يخفى في هذا الصدد، وهي: علم المعلومات، والإحصاء، وعلم الكمبيوتر، أما رابعها فهو قطاع عام يمثل مجالات معرفية مختلفة. ويتضح من الشكل أن المجالات الثلاثة تبدو أكثر تمايزًا مما هي عليه في الواقع العملي؛ حيث نجد ـ على سبيل المثال ـ أن ربع أعضاء هيئة التدريس في مدرسة المعلومات في تشابل هيل Chapel Hill حاصلون على درجة الدكتوراه في علوم الحاسب، فيما حصل الآخرون على درجاتهم العلمية في مجالات مثل: علم الاجتماع، والتربية، والفيزياء، والإدارة، بالإضافة إلى مجال علم المكتبات والمعلومات.

ولتوضيح ما يسهم به كل قطاع من القطاعات الأربعة، سنشرحها بإيجاز شديد، مع مزيد من التركيز على علم المعلومات.



الشكل (1)أسس علم البيانات

من المؤكد أن قطاع مجال المعرفة سيتسع لمعظم العلماء المشتغلين بممارسة علم البيانات، ولا شك أن كل مجال معرفي سيؤثر بقوة في أنواع البيانات التي يتم تداولها ـ وفي القرارات التي يتم اتخاذها بناءً على التحليلات، وسوف

يقوم الكيميائيون وعلماء السياسة والأطباء والصحفيون والمحاسبون وغيرهم من العلماء الأكاديميين في جميع المجالات، بتبني ثم تكييف الأدوات والأساليب التي ابتكرها علماء البيانات، وسيرغب العديد من هؤلاء في التخصص في علوم البيانات كلٌّ في مجاله؛ ذلك أن المجال التخصصي يحدد ـ بشكل دقيق ـ البيانات التي يتم جمعها، وما وراء البيانات الأكثر أهمية لفهم تلك البيانات، كذلك كيفية تحديد جودة البيانات، وقيمة النتائج التحليلية المستخلصة منها. وبالتالي سيتعين على أي برنامج تدريبي في علم البيانات، النظر في كيفية دمج خبرات المجال التخصصي في المناهج الدراسية، والمشروعات الأساسية، والأنشطة البحثية للطلاب.

لقد ظهر علم الإحصاء أول مرة بوصفه مجالًا فرعيًّا تطبيقيًّا للرياضيات، ثم وضعت على مدار المائة والخمسين عامًا الماضية مبادئ الاحتمالات بهدف استحداث نظريات أخذ العينات، والتقدير وخطأ التقدير، والدلالة الإحصائية، كما تأسس أول أقسام الإحصاء، وأول مجلة متخصصة في مطلع القرن العشرين الميلادي المنصرم، ويقوم الإحصائيون اليوم باستحداث أساليب جديدة لتلخيص مجموعات البيانات (datasets)، واستخراج وتقييم الاستنتاجات من البيانات، وابتكار نماذج جديدة لتحليل المعنى، وإجراء التنبؤات المبنية على البيانات. وعليه يمكن القول إن النظرية الإحصائية وأساليبها تسهم في علم البيانات من خلال اتخاذ قرارات حول البيانات التي يجب جمعها أو تضمينها في التحليل (مثل: أخذ العينات)، وتقديم أساليب إحصائية ـ لربما تكون آلية، وتقييم فاعلية النتائج (مثل : تقدير الخطأ).

أما علم الحاسب فقد ظهر أول مرة في النصف الأول من القرن العشرين الميلادي بعد أن صمم الرياضيون والفيزيائيون آلات يمكنها حساب البيانات والمقارنة بينها بسرعات تفوق قدرات البشر، وأدت الأفكار الرئيسية المنبثقة من هذا العلم، مثل تخزين تعليمات التشغيل (البرمجيات)، وكذلك البيانات وتقدير التعقيدات الحسابية، إلى ظهور النظرية الحاسوبية التي ازدهرت في منتصف القرن العشرين الميلادي في مجال جديد أعطى للأقسام الأكاديمية والمجلات العلمية والجمعيات المهنية ذلك الشكل الذي هي عليه اليوم. ويقوم علماء الحاسبات بإنشاء أو تطويع خوارزميات رياضية وإحصائية تيسر معالجة البيانات عالية السرعة وكبيرة الحجم. وعلى نحو متزايد يعمل علماء الحاسبات مع خبراء من مجالات متعددة لنمذجة توزيعات البيانات وأنماطها، وديناميكيتها باستخدام أساليب تجريبية (مثل تعلم الآلة Machine Learning)، كما يسهم علم الحاسبات في علم البيانات من خلال توفير أساليب للآلات وللخوارزميات التي يمكن لخبراء المجالات تطبيقها في طرح الأسئلة والإجابة عنها.

أما المجال الذي نعرفه باسم "علم المعلومات" فيعود تاريخه إلى نهاية القرن التاسع عشر عندما أسس بول أوتليه وهنري لافونتين Paul Otlet and Henri La Fontaine حركة التوثيق في أوروبا، وهي حركة حولت تركيز

اختصاصيي المكتبات من جمع المعرفة من مصادرها التاريخية إلى الإدارة المنهجية لتلك المصادر، وعكست إمكانية تطوير ممارسات الإدارة العلمية (نظرية تيلور للإدارة العلمية Taylorism) في المجال. ثم بدأ راماميريتا رانجاناثان Shiyali Ramamrita Ranganathan وآخرون في استخدام مصطلح "علم المكتبات" ـ بوصفه مصطلحًا يدل على الأساليب المنظمة في إدارة وتعليم المكتبات. وقد أدى تطور نظرية المعلومات على أيدي علماء مثل "آلان تورينج Alan Turing" ، و "نوربرت وينر Norbert Wiener " إلى ظهور أجهزة حاسبات تناظرية ورقمية في أربعينيات القرن العشرين الميلادي، ثم دفعت الحرب العالمية الثانية عجلة التقدم في مجال الاتصالات، وجمع المعلومات الاستخبارية، وإدارتها، وتفسيرها (تشفير المعلومات Cryptography، والنمذجة الإحصائية Statistical Modeling التي مهدت لتقنيات تنقيب البيانات Data Mining)، هذا بالإضافة إلى أساليب التحكم أو ضبط المعلومات (بحوث العمليات) لإدارة أعداد كبيرة من البشر ومستلزماتهم (الجنود والذخائر والإمدادات) التي تتناقل في جميع أنحاء الكوكب بأكمله. كما مهدت رؤية فانيفار بوش Vannevar Bush لجهاز الميمكس Memex (1945) الذي بنيت فكرته على تقنية الميكروفورم Microform، لظهور النصوص التشعبية (hypertext) وشبكة الويب. كما يعد العمل الريادي الذي قدمه كلود شانون Claude Shannon حول المعلومات في عام 1948 بمثابة القاعدة التي قامت عليها نظرية المعلومات؛ فمبوجب هذا العمل (Shannon & Weaver, 1948) عرفت المعلومات بأنها مقدار عدم اليقين Uncertainty الذي يتناقص مع وجود رسالة، كما عرفت وحدة البت bit (الوحدة الحسابية الثنائية Binary) باعتبارها الوحدة الأساسية للمعلومات، وكذلك القدرة القصوى لسعة الحزمة الموجبة Bandwidth لقنوات الاتصال.

ومجال علم المعلومات يبحث اليوم في معماريات وأنطولوجيات (تصميم، وتوصيف) مجموعات البيانات؛ وفي عمليات معالجة وإدارة الأصول الرقمية، وفي أخلاقيات وسياسات المعلومات، ونشوء وتدفق وحفظ المعرفة، والتفاعل بين المعلومات والإنسان، هذا فضلًا عن مجموعة متنوعة من تطبيقات مبادئ وممارسات علم المعلومات في مجالات الصحة (المعلوماتية الصحية Health Informatics)، والإنسانيات (الإنسانيات الرقمية Digital Humanities)، والتجارة (المعلوماتية المالية Financial Informatics)، والقانون والحكومة (المعلوماتية القانونية Legal Informatics)، والبيئة (المعلوماتية البيئية Environmental Informatics).

وكما أشرنا في الفقرة السابقة، يهتم علماء المعلومات بالدورة الكاملة لحياة البيانات، وبالقضايا الاجتماعية – الثقافية المرتبطة بجمع البيانات واستخدامها. ومن أمثلة ذلك:

- الضوابط الأخلاقية والقانونية المرتبطة بجمع البيانات (مثل: الموافقة المسبقة Informed Consent، والخصوصية، واللوائح القانونية).
- التقييم، وجودة البيانات، وتنقية البيانات.
- إضافة ما وراء البيانات للوثائق (مثل: وحدات القياس، وضوابط عمل أجهزة الاستشعار والنقل، وضبط المصطلحات أو المفردات)، وذلك بهدف ضمان كفاءة الآلات، وكفاءة تفسير البشر عند التعامل مع تلك الوثائق.
- اختزان وحفظ البيانات (مثل سياسات نسخ البيانات، وسياسات التثبت من صحتها)، وما يتعلق بذلك من معالجة واستخدام (مثل توثيق تنقية البيانات Data Cleaning، وخوارزميات التحليل، والخرائط الكاملة لسير أو تدفق البيانات Workflows).
- تقييم الاستنتاجات المبنية على الاستفادة من البيانات وتحليلها، وجعل البيانات وخرائط التدفق قابلة للبحث وإعادة الاستخدام.

في ضوء ما سبق يمكن القول بأن علم البيانات هو بمثابة مجموعة فرعية من علم المعلومات، وعلى الرغم من وجود برامج تدريبية لعلم البيانات في مدارس المعلومات، فإنه من الأفضل ـ من الناحية الاستراتيجية ـ النظر إلى علم المعلومات على أنه مكون أساسي من مكونات علم البيانات بحيث يستطيع هذا المجال الوليد أن يستفيد من تنوع وجهات النظر التي تتمخض عن التعاون بين تخصصات متعددة. ويمكن لبرامج علم المعلومات المساهمة بشكل كبير وفعّال في قيادة الطلاب للإشكالات الاجتماعية والثقافية التي أشرنا إليها من قبل. وبالإضافة إلى ذلك، تستطيع مدارس المعلومات التي لديها هيئة تدريس ممن يمتلكون مهارات عالية في مجال التقنية أن تساعد الطلاب على تطوير مهاراتهم الفنية في إدارة البيانات الموزعة أو الشبكية (مثل تنفيذ حلول تعتمد على نظم إدارة قواعد البيانات مثل NoSQL، ووضع قواعد مبنية على سياسات خاصة بنسخ البيانات، وأمن المعلومات، وإمكانات الوصول إليها)؛ وتطبيق عمليات اللغة الطبيعية، وتنقيب البيانات، ومجموعات برمجيات تعلم الآلة لأغراض التحليل، كذلك عمل نصوص Scripts مخصصة للمهام الإدارية، وإعداد التقارير (مثل استخدام لغة البرمجة المعروفة ب R أو بايثون Python)؛ ووضع نظم الاستعلام المناسبة، ونظم إعداد التقارير ـ ونظم التمثيل التصويري للبيانات والمخرجات؛ كذلك القيام بتدقيق ماوراء البيانات من خلال استخدام معايير متخصصة، أو باستخدام الأنطولوجيات (مثل: مجموعة أدوات HIVE[6]). ومع استمرار تطور مجال علم البيانات، سوف تكون هناك حاجة أيضًا لتطوير مهارات محددة يحتاج إليها علماء البيانات والمشتغلون بهذا العلم؛ لذلك يجب أن تسهم برامج علم المعلومات ـ بشكل نشط ـ في الفرق متعددة التخصصات التي ستشكل هذا المجال.

ومن نافلة القول أن المعلومات في حد ذاتها هي ظاهرة محورية أو أساسية تحظى بالاهتمام وقائمة بذاتها، ومدارس المعلومات تمثل كيانات هادفة تتسم بالموضوعية، ولها أهمية بالغة في تعليم الباحثين والممارسين الذين يعملون في مجموعة واسعة من المؤسسات، أما علم البيانات ما هو إلا مجال واحد ناشئ سوف يستفيد ـ بلا شك ـ من انخراطه بمدراس المعلومات.

●●●

المصادر

Borgman, C. (2015). Big data, little data, no data: Scholarship in the networked world. Cambridge, MA: MIT Press. Bush, V. (1945). As we may think. The Atlantic, 7. Retrieved from http://www.theatlantic.com/ magazine/archive/1945/07/as-we-may-think/303881/.

Johnson, J. (2016). The question of information justice. Communications of the ACM, 59(3), 27–29. Hey, T., Tansley, S., & Tolle, K. (2009). The fourth paradigm: Data-intensive scientific discovery. Redmond, WA: Microsoft Corporation. Retrieved from http://research.microsoft.com/en-us/ collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf.

McAfee, A. &Brynjolfsson, E. (2012). Big data: The management revolution. Harvard Business Review, 90(10), 60–68.

Shannon, C.E., & Weaver, W. (1948). The mathematical theory of communication. Bell System Technical Journal, 27 (3), 379–423. Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. In Proceedings of the 8th International AAAI Conference on Weblogs and Social Media (ICWSM '14) (pp. 505–514). Palo Alto, CA: AAAI Press. Zikopoulos, P., & Eaton, C. (2011). Understanding big data: Analytics for enterprise class hadoop and streaming data (1st ed.). New York, NY: McGraw-Hill Osborne Media.

●●●

الحواشي السفلية

1 العنوان الأصلي للمقالة هو: Information Science Roles in the EmergingField of Data Science. وقد نشرت المقالة في مجلة: Journal of Data and Information Science, JDIS Vol. 1 No. 2, 2016 pp 1–6

2 جارى مارشيونيني Gary Marchionini: هو أستاذ وعميد كلية علوم المعلومات والمكتبات بجامعة نورث كارولينا Universityof North Carolina- في تشابل هيل. نشر مارشيونيني أكثر من 200 مقالة وفصل وتقرير في مجموعة متنوعة من الكتب والمجلات. وقد عمل مارشيونيني رئيس تحرير لمجلة ACM المتخصصة في نظم المعلومات - خلال الفترة (2002-2008)، كما عمل محررًا لسلسلة محاضرات Morgan-Claypool المتخصصة في المعلومات من حيث: المفاهيم، ونظم الاسترجاع، والخدمات. كذلك شغل مارشيونيني منصب رئيس برنامج ACM SIGIR في عام 2005 , وبرنامج ACM / IEEE JCDL - في عام 2002 ، كما ترأس ACM DL 96 و JCDL 2006. فضلًا عن ذلك كان ولا يزال عضوًا في لجان تحرير العديد من المجلات العلمية المتخصصة، كما شغل منصب رئيس جمعية علم المعلومات والتقنية. حصل مارشيونيني بحكم تميزه على تمويل من هيئات شهيرة مثل المؤسسة الوطنية للعلوم في الولايات المتحدة الأمريكية، ومن شركات Microsoft و Google و.IBM ، ومن أهم اهتماماته ومشروعاته ما يتصل بمجال الواجهات التي تدعم البحث عن المعلومات واسترجاع المعلومات، قابلية استخدام السجلات الصحية الشخصية، والتمثيل البديل للوثائق الإلكترونية، واستراتيجيات تصفح الوسائط المتعددة، والمكتبات الرقمية، وتقييم الوسائط التفاعلية، وبخاصة ما يعنى منها بمجالي التعليم والتعلم. وللمزيد يمكن مراجعة الموقع الشخصي للمؤلف على الرابط التالي:

https://sils.unc.edu/people/faculty/profiles/Gary-Marchionini

3 مع زيادة الحاجة لزيادة حجم البيانات والتأكد من سلامتها وعدم خسارة جزء منها، كان لابد من وجود نسخ مكررة فعالة ومتزامنة بحيث عند وجود أي مشكلات في أي بيانات يمكن استبدالها، وكذلك توزيع الحمل على أكثر من جهاز وليس جهاز واحد، ومن هنا ظهرت قواعد البيانات أمثال Oracle وMysql وMS Sql Server ، وغيرها بتبني فكرة أن تكون قواعد البيانات تعمل بنظام موزع ومتواز، وهو ما يعرف Parallel DBMS لتخزين البيانات [المترجم].

4 من المبادئ التي نشرتها جوجل ليكون نواة للكثير من البرمجيات التي تعمل على حل المشكلات الضخمة، وهذا المبدأ يقوم على تقسيم المشكلة الكبيرة إلى عدة أجزاء بحيث يمكن معالجة كل جزء على حدة، وبعد ذلك يتم تجميع هذه الحلول في حل واحد ليكون حل المشكلة الكبيرة، وبهذه الطريقة أصبحت التكلفة أقل [المترجم].

5 هي مفهوم يقصد به وجود كم كبير من البيانات المنظمة وغير المنظمة معًا.[المترجم].

6 من برمجيات أباتشي لتيسيير التعامل مع البيانات الضخمة [المترجم].

●●●