



البقاة للمعلوماتية

مجلة دورية تصدرها الجمعية العالمية السورية للمعلوماتية

السنة العشرون - العدد الواحد والخمسون - آذار "مارس" / حزيران "يونية" 2016

ملف العدد: تحديات في المعطيات الكبيرة

الحوسبة بمقياس إكسا، والمعطيات الكبيرة

المعطيات الكبيرة وتحدياتها التقنية

الأبحاث الأخرى

الأمن والخصوصية لأنظمة الحقيقة المزيدة

المخاطر والأساطير المتعلقة بالحوسبة السحابية والتخزين السحابي

يسبر الباحثون الأمن بواسطة التعقيم

أنظمة إسقاط المصوّرات التجسيمية توفر حياةً أبدية

قراءة الأدمغة

استعمال معطيات المريض لإجراء معالجات سرطان مُخصّصة

تعليم الحواسيب الخدع البصرية



البَقَاةُ الْمَعْلُومَاتِيَّةُ

مَجَلَّةٌ دَوْرِيَّةٌ تُصَدِّرُهَا الْجَمْعِيَّةُ الْعَامِيَّةُ السُّورِيَّةُ لِلْمَعْلُومَاتِيَّةِ

السنة العشرون - العدد الواحد والخمسون - آذار "مارس" / حزيران "يونية" 2016

رئيس التحرير:

الدكتور موفق دعبول

هيئة التحرير:

الدكتور سعد الله آغا القلعة

الدكتور سامح جزماتي

الدكتور نزار الحافظ

الدكتور راكان رزوق

الدكتور حسان ريشة

الدكتور عماد الصابوني

الدكتورة ندى غنيم

الدكتور منصور فرح

الدكتور محمد مراياتي

أمانة التحرير:

هيفاء باكير

للمراسلات:

الجمعية العلمية السورية للمعلوماتية

مجلة الثقافة المعلوماتية

دمشق - الجمارك - بجانب وزارة التعليم العالي

ص.ب. 33492 - سورية

هاتف: **2116155**

فكس: **2127998**

بريد إلكتروني: nzhafez@scs-net.org

تنسيق المقالات وإعداد المجلة للطباعة: **الدكتور نزار الحافظ**

الإخراج الفني والمعالجة: **مركز الفوال للفرز والمونتاج الإلكتروني**

عززي القاري،

اخترنا لهذا العدد بحثين تحت العنوان **تحديات في المعطيات الكبيرة**، هما: " الحوسبة بمقياس إكسا، والمعطيات الكبيرة "، " المعطيات الكبيرة وتحدياتها التقنية ".

يتناول **البحث الأول** (الحوسبة بمقياس إكسا، والمعطيات الكبيرة) الحوسبة المتقدمة اللازمة للمتكمين للاكتشافات العلمية والابتكارات الهندسية. بعد تمهيد لتاريخ الحوسبة المتقدمة، يستعرض المقال بعض التحديات التقنية، والارتباط المتبادل بين النمذجة الحاسوبية وتحليلات المعطيات الكبيرة، والنظام البيئي الشمولي، وتتألف الشركات على صدارة الحوسبة المتقدمة.

يعطي **البحث الثاني** (المعطيات الكبيرة وتحدياتها التقنية) مقدمة إلى المعطيات الكبيرة، ويناقش مراحل دورة حياة المعطيات الكبيرة من البداية حتى النهاية (تحصيل المعطيات، استخراج المعلومات وتنقيتها، مكاملة المعطيات وجمعها وتمثيلها، النمذجة والتحليل)، والتحديات الخاصة بكل مرحلة (عدم التجانس، عدم الاكتمال وعدم التوافق، تصدع المقياس، التوقيت المناسب، الخصوصية وملكية المعطيات). ويقدم دراسة حالة تُبرز مثالاً للقضايا التي تنشأ في المراحل المختلفة، مستعرضاً التحديات المُتشابهة المتعلقة بها.

واخترنا أيضاً لهذا العدد سبعة بحوث متنوعة المواضيع مثيرة للاهتمام هي على الترتيب: " الأمن والخصوصية لأنظمة الحقيقة المزيدة "، " المخاطر والأساطير المتعلقة بالحوسبة السحابية والتخزين السحابي "، " يسبر الباحثون الأمن بواسطة التعنيم "، " أنظمة إسقاط المصورات التجسيمية توفر حياة أبدية "، " قراءة الأدمغة "، " استعمال معطيات المريض لإجراء معالجات سرطان مُخصّصة "، " تعليم الحواسيب الخدع البصرية ".

يلقي **أول هذه البحوث** (الأمن والخصوصية لأنظمة الحقيقة المزيدة) نظرة على فضاء الحقيقة المزيدة، مع الاهتمام بالتطبيقات المباشرة لأنظمة الحقيقة المزيدة والتقانات اللازمة لدعم هذه التطبيقات. ويستكشف التحديات الجديدة في مجال الأمن والخصوصية، التي تولدها تقانات الحقيقة المزيدة، والتوجهات الدفاعية، والتطبيقات الجديدة لأنظمة الحقيقة المزيدة للتصدي لمسائل الأمن والخصوصية المعروفة.

يدرس **ثانيها** (المخاطر والأساطير المتعلقة بالحوسبة السحابية والتخزين السحابي) خدمات السحاب وميزاتها، ويناقش أنواعاً جديدة من المخاطر التي تبديها الخدمات السحابية، أهمها الوثوق بمقدّم غير موثوق به للتخزين البعيد. ويشير المقال إلى حلول بناءً للتعامل مع تلك المخاطر.

يتحدث **ثالثها** (يسبر الباحثون الأمن بواسطة التعنيم) عن التعنيم الذي يجعل البرامج الحاسوبية غير مفهومة بحيث لا يمكن للمهاجمين فهم كيفية عملها. ويستعرض مفاهيم تعمية مثل الصندوق الأسود الافتراضي، والتعنيم غير المميز الذي يمكن أن يكون أساساً لتقانات جديدة للتعمية.

يستقضي **رابعها** (أنظمة إسقاط المصورات التجسيمية توفر حياة أبدية) تقنية التصوير التجسيمي التي تسجل مشهداً ثلاثي الأبعاد على صورة ذات بعدين. ويبين المقال كيفية استعمال تقانة Eyeliner لإعادة ممثلين أو فنانيين متوفّين إلى الحياة مجدداً.

يسلط **خامسها** (قراءة الأدمغة) الضوء على الجديد في موضوع قراءة العقل الذي أصبح من مجالات العلم الجادة، ويبرز فكرة إنشاء واجهات بين الدماغ البشري والآلة تمكّن من التعامل مع الحواسيب والآلات عن طريق الأفكار. ويستعرض نشاطات الباحثين الراهنة في مجال قراءة أفكار الدماغ، ومشكلة حماية خصوصية الأفكار.

يلقي **سادسها** (استعمال معطيات المريض لإجراء معالجات سرطان مُخصّصة) نظرةً على أهمية إنشاء قواعد معطيات ضخمة آنية لمعطيات المرضى المصابين بأمراض خطيرة كالسرطان، في تعزيز ما يسمّى الطب المخصّص (الذي يقدم علاجًا أشدّ فعالية للأمراض الخطيرة). ويستكشف فرص نفاذ الأطباء الممارسين المباشر إلى المعطيات الطبية المناسبة المخزنة في قواعد المعطيات تلك.

أخيراً، يتأمل **سابعها** (تعليم الحواسيب الخدع البصرية) كيفية تعرّف الصور البصرية بالحاسوب، ويركّز على رؤية الانطباعات البصرية الخادعة وأساليب تحسين آلية الرؤية هذه بالحاسوب. ويستعرض الجهود الجارية لتدريب الحاسوب (أو الربوط) على تعرّف الصور البصرية.

هذا وقد أرجأنا قائمة المصطلحات (إنكليزي - عربي) الواردة في هذه البحوث التسعة إلى العدد القادم من المجلة.

أخي القارئ،

أصدرت الجمعية العلمية السورية للمعلوماتية كتاب " البرمجة المتوازية - تقنياتها وتطبيقاتها باستعمال محطات عمل شبكية وحواسيب متوازية"، وهو الكتاب العاشر ضمن سلسلة الكتب التي تصدرها الجمعية.

في الختام، نتمنى أن تقدم مواضيع هذا العدد الفائدة المرجوة، ونأمل أن نتواصل معنا بإرسال ملاحظاتك ومقترحاتك إلينا ...

وإلى اللقاء معك في العدد القادم.

رئيس التحرير

الدكتور موفق دعبول

المحتويات

ملف العدد: تحديات في المعطيات الكبيرة

- 9 الحوسبة بمقياس إكساء، والمعطيات الكبيرة
- 31 المعطيات الكبيرة وتحدياتها التقنية

الأبحاث الأخرى

- 45 الأمن والخصوصية لأنظمة الحقيقة المزيدة
- 61 المخاطر والأساطير المتعلقة بالحوسبة السحابية والتخزين السحابي
- 67 يسبر الباحثون الأمن بواسطة التعقيم
- 71 أنظمة إسقاط المصوّرات التجسيمية توفر حياةً أبدية
- 77 قراءة الأدمغة
- 83 استعمال معطيات المريض لإجراء معالجات سرطان مُخصّصة
- 89 تعليم الحواسيب الخدع البصرية

الحوسبة بمقياس إكسا، والمعطيات الكبيرة

EXASCALE COMPUTING AND BIG DATA*

Daniel A. Reed, Jack Dongarra

ترجمة: د. نزار الحافظ¹
مراجعة: د. مكي الحسني

تتطلب الاكتشافات العلمية والابتكار الهندسي توحيد جانبيين مفصولين تقليدياً: الحوسبة العالية الأداء وتحليلات المعطيات الكبيرة.

قبل قرنين تقريباً، كتب الكيميائي الإنكليزي همفري ديفي Humphrey Davy: "لا شيء يميل كثيراً لتطوير المعرفة كما يفعله تطبيق أداة جديدة. إن قدرات الرجال الفكرية الأصيلة في الأوقات المختلفة ليست إلى حد بعيد هي الأسباب التي أدت إلى النجاحات المختلفة لجهودهم، بل الطبيعة الخاصة من الوسائل والموارد الاصطناعية التي يملكونها." إن ملاحظة

ديفي، أن الميزة هي لأولئك الذين لديهم أقوى الأدوات العلمية، ليست أقل صحة اليوم. ففي عام 2013، حصل مارتن كارلس وميشال ليفيت وأرييه وارشل Martin Karplus, Michael Levitt, and Arie Warshel على جائزة نوبل في الكيمياء على أعمالهم في النمذجة الحسابية. إذ قالت لجنة نوبل: أصبحت النماذج الحاسوبية التي تجسد الحياة الحقيقية حاسمة في معظم التقدم الذي أحرز في الكيمياء اليوم، [17]، و "تكشف الحواسيب النقاب عن عمليات كيميائية، مثل تنقية أبخرة العوادم بواسطة حافز أو عملية التمثيل الضوئي في الأوراق الخضراء".

وسواء كنا نَصِف مزايا مسرعات الجسيمات ذات

نقاط مفتاحية

- تباينت أدوات الحوسبة العالية الأداء وثقافتها وتحليلات المعطيات الكبيرة، على حساب كل منها؛ وتوحيدها لا بد منه لمواجهة طيف من المجالات البحثية الكبرى.
- تتطلب تحديات المقياس منا زيادة قدرتنا على نقل المعطيات أو حساب المهام المعقدة عليها، أو تخزين جزء كبير منها؛ وثمة حاجة إلى مناهج جديدة لمواجهة هذه التحديات.
- تتطلب الطبيعة الدولية للعلوم المزيد من تطوير بنى حاسوبية متقدمة ومعايير عالمية لمعالجة المعطيات، على حين تُعقد المنافسة الدولية انفتاح الإجراءات العلمية.

* نُشر هذا البحث في مجلة *Communications of the ACM*، المجلد 58، العدد 7، تموز (يوليه) 2015، الصفحات 56 – 68.

¹ مدير بحوث في المعهد العالي للعلوم التطبيقية والتكنولوجيا.

الطاقة العالية (مثل مصادم الجسيمات الضخم Large Hadron Collider واكتشاف بوزون هيغز Higgs boson في عام 2013)، أو أدوات علم الفلك القوية (مثل مقراب (تلسكوب) الفضاء هابل Hubble ، الذي أسفر عن نظرة ثاقبة إلى توسع الكون والطاقة الخفية)، أو مُسلسلات DNA (DNA sequencers) ذات الإنتاجية العالية واكتشاف إيكولوجية (بيئة) المجينات الفائقة metagenomics، فإن الأدوات العلمية المتزايدة دائماً، تدفع باستمرار عجلة تطور المعرفة. وإن كل أداة علمية من هذا القبيل، وكذلك مجموعة من الأدوات الأخرى، تعتمد اعتماداً حاسماً على حوسبة التحكم في المُحسّنات، ومعالجة المعطيات، والتعاون الدولي، وإمكان النفاذ.

مع ذلك، فالحوسبة هي أكثر بكثير من مجرد مُعزّز للعلم. إن النمذجة الحاسوبية وتحليلات المعطيات، خلافاً لغيرها من الأدوات التي تقتصر على مجالات علمية معيّنة، تُطبّق على جميع مجالات العلوم والهندسة، ذلك أنها تبعث الحياة في الرياضيات الكامنة وراء النماذج العلمية. وهي تمكّن الباحثين من فهم توقّعات مختلفة بعض الشيء، كما أنها تعطي التجارب شكلاً أكثر فعالية. وتساعد أيضاً على التقاط وتحليل سيل من المعطيات التجريبية التي ينتجها جيل جديد من الأجهزة العلمية الذي تحقّق بفضل التقدم في مجال الحوسبة والإلكترونيات الدقيقة microelectronics.

يمكن للنمذجة الحاسوبية أن تُضيء دقائق النماذج الرياضية المعقدة وتدفع عجلة تطور العلوم والهندسة، حيث الوقت أو التكلفة أو السلامة تُحوّل دون التقييم التجريبي وحده. لقد وفّرت النماذج الحاسوبية الخاصة بالظواهر الفيزيائية الفلكية - على مقاييس زمنية ومكانية متنوعة بتنوع تشكيل نظام الكواكب - والديناميكية النجمية، وسلوك الثقب الأسود، وتشكيل المجرات، وتفاعل المادة المظلمة الباريونية العتيدة، رؤى جديدة في النظريات والمعطيات التجريبية المستكملة. وكانت النماذج المناخية المتطورة التي التقطت آثار غازات الدفيئة greenhouse gases، وإزالة الغابات، والتغيرات الأخرى على الكواكب، المفتاح لفهم الآثار المترتبة على السلوك البشري في تغير الطقس والمناخ.

تتيح العلوم والهندسة الحاسوبية أيضاً التصميم المتعدد التخصصات والاستمثال، وهذا ما يخفض مدة النمذجة الأولية وتكلفتها. وقد مكّنت المحاكاة المتطورة شركة Cummins من بناء محركات ديزل أفضل وأسرع وبتكلفة أقل، وشركة غودبير Goodyear من تصميم إطارات أكثر أماناً بسرعة أكبر بكثير، وشركة بوينغ من بناء طائرات أعلى كفاءةً في استهلاك الوقود، وشركة بروكتر وغامبل Procter & Gamble من صنع مواد أفضل للمنتجات المنزلية.

وبالمثل، اعتُبرت "المعطيات الكبيرة"، والتعلم الآلي، وتحليل المعطيات التنبؤية النموذج الرابع للعلوم، [12]، وهذا ما يسمح للباحثين باستخراج الأفكار من كل من الأدوات العلمية والمحاكاة الحاسوبية. أعطى التعلم الآلي رؤى جديدة للمخاطر الصحية المحتملة وانتشار الأمراض، عن طريق تحليل الشبكات الاجتماعية، واستعلامات البحث في الوب، ومعطيات المستشفى. بل هو أيضاً المفتاح لتحديد الأحداث وارتباطاتها في مجالات متنوعة مثل فيزياء الطاقة العالية والبيولوجيا الجزيئية.

كما هو الحال مع الأجيال المتعاقبة من الأدوات العلمية الأخرى الواسعة النطاق، يوفّر كل جيل جديد من الحوسبة المتقدمة قدراتٍ جديدة، جنباً إلى جنب مع تحديات التصميم التقني والتوازنات الاقتصادية. وبوجه عام، تنمو قدرات توليد المعطيات في معظم المجالات العلمية بسرعة أكبر من قدرات الحوسبة، وهذا ما يجعل هذه المجالات كثيفة المعطيات data-intensive [23]. ترتبط الحواسيب العالية الأداء ونظم المعطيات الكبيرة ارتباطاً وثيقاً بالنظام البيئي للحوسبة الأوسع ويتصميماته وأسواقه. وهي أيضاً تدعم احتياجات الأمن القومي والقدرة التنافسية الاقتصادية بالطرق التي تميزها عن معظم الأدوات العلمية الأخرى.

هذا النموذج "ذو الاستعمال المزدوج"، جنباً إلى جنب مع ارتفاع تكاليف نُظُم متزايدة الكِبَر دائماً للحوسبة وتحليل المعطيات، إلى جانب مجموعة من تحديات جديدة للتصميم على نطاق واسع، يُثير تساؤلاتٍ جديدة حول الاستثمار والأولويات والتصميم ونماذج المشتريات، إضافة إلى التعاون والتنافس العالمي، في بحوث الحوسبة المتقدمة. يتناول هذا المقال بعضَ هذه التحديات التقنية، والترابطَ بين النمذجة الحسابية وتحليلات المعطيات، والنظامَ البيئي العالمي والتنافس على القيادة في مجال الحوسبة المتقدمة. ونبدأ بالتمهيد لتاريخ الحوسبة المتقدمة.

النظم البيئية للحوسبة المتقدمة

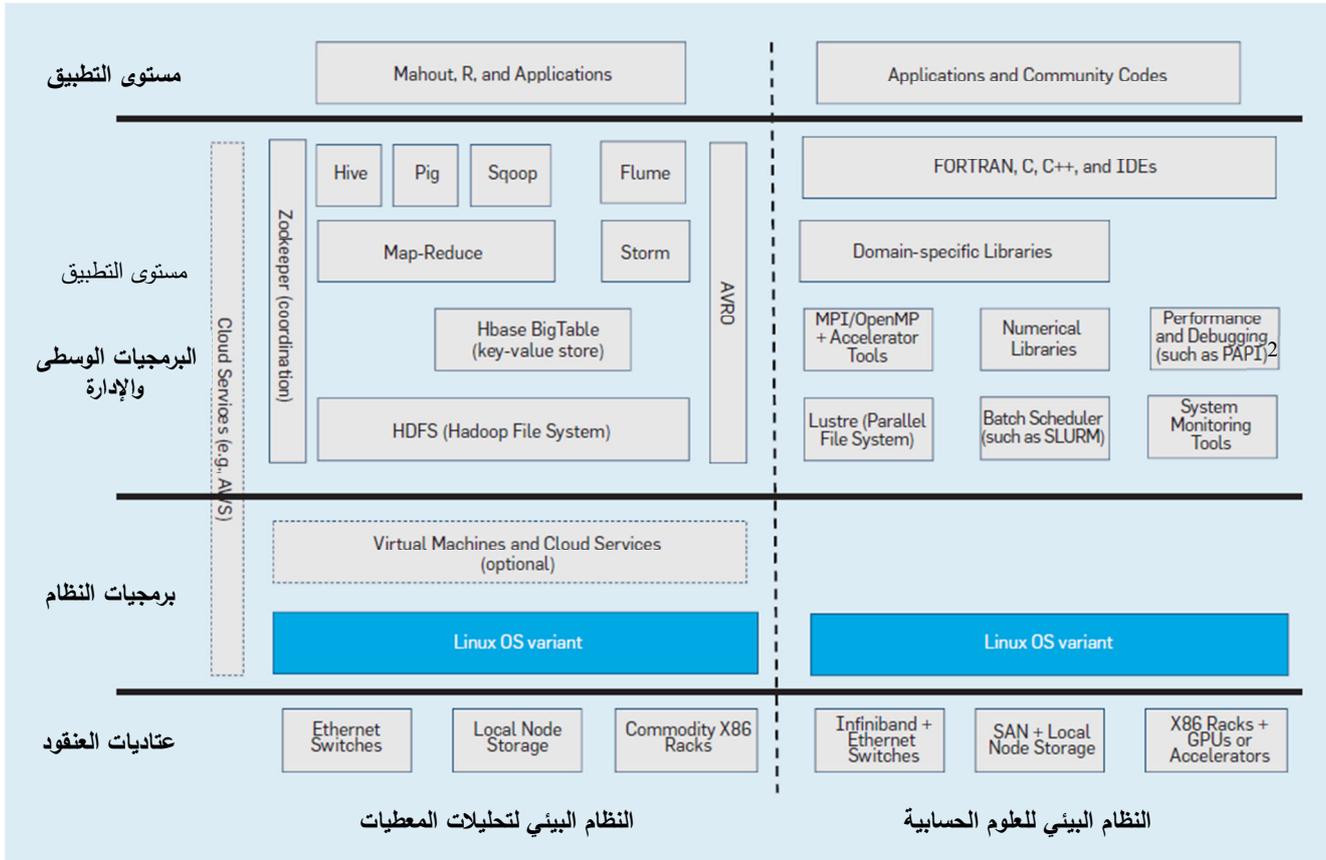
بالتعريف، يتضمن نظام الحوسبة المتقدمة العتادَ، والبرمجيات، والخوارزميات اللازمة لتقديم المستوى الأعلى من الإمكانيات في أي وقت من الأوقات. وكما في الشكل 1، تتشارك النظم البيئية للحوسبة وتحليلات المعطيات في بعض الصفات، وبالأخص في الاعتماد على البرمجيات المفتوحة المصدر والنظام البيئي للعتاد x86. ومع ذلك، فإنها تختلف على نحو ملحوظ في اهتماماتها ومناهجها التقنية. ولما كان البحث العلمي يعتمد على نحو متزايد على كلا الحوسبة العالية السرعة وتحليلات المعطيات، فإن التقارب المحتمل للتشغيل البيئي وتوسيع النطاق لهذين النظامين البيئيين هو أمر حاسم في المستقبل.

الحوسبة العلمية. في الثمانينيات، هيمنت الحوسبة الفائقة للمتجهات vector supercomputing على الحوسبة العالية الأداء، على النحو المنصوص عليه في نُظُم صممها الراحل سيمور كراي Seymour Cray بإسمه. وشهدت التسعينيات صعودَ المعالجة المتوازية الغزيرة (massively parallel processing (MPP)، ومتعدد المعالجات بذاكرة مشتركة (SMPS) shared memory multiprocessors التي بُنيت في الآلات المفكرة Thinking Machines، وسيليكون جرافيكس Silicon Graphics، وغيرها. في المقابل، هيمنت عناقيدٌ سلعية (معالجات x86 من إنتل و AMD)، ومعالجات ذات أغراض محددة (مثل BlueGene من IBM)، في العقد السابق.

واليوم، يُضاف إلى هذه العناقيد مسرعاتٌ حسابية في شكل معالجات مساعدة coprocessors من إنتل ووحدات معالجة بيانية (GPUs) من إنفيديا Nvidia؛ وهي أيضاً تتضمن وصلاتٍ بينية interconnects عالية السرعة منخفضة التلبيث (مثل InfiniBand إنفينيباند). وتُستعمل شبكات مناطق التخزين (SANs) لتخزين المعطيات الدائمة، مع أقراص محلية على كل عقدة لا تُستعمل إلا للملفات المؤقتة. وجرى استمثال هذا النظام البيئي العتادي لأجل الأداء أولاً، لا لأجل أقل تكلفة ممكنة.

ويوفر لينكس، فوق عتاد العنقود، تجهيزاتٍ نظام، مضافاً إليها نُظُم ملفات متوازية (مثل أستر Lustre) ومجدولات دُفعية (مثل PBS و SLURM) لإدارة الوظائف المتوازية. ويُستعمل MPI و OpenMP للتوازي بين العقد وداخلها (internode and intranode parallelism)، مُضافاً إليها مكتبات وأدوات (مثل CUDA و OpenCL) لاستعمال المعالج المساعد. إن المكتبات الرقمية (مثل LAPACK و PETSc) والمكتبات الخاصة بمجالات محددة، تُكمل كدسة البرمجيات. وعادة يجري تطوير التطبيقات بلغات C, C++, FORTRAN.

تحليلات المعطيات. قبل بضع سنوات، لم تحو أكبر نظم تخزين المعطيات سوى عدد قليل من التيرابايتات terabytes من الخزن الثانوي على القرص، تُدعمها مكتباتُ الأشرطة المؤتمتة. أما اليوم، فإن نظم الحوسبة السحابية



الشكل 1. مقارنة بين تحليلات المعطيات والنظام البيئي للحوسبة.

التجارية والبحثية، كلٌّ منها تحوي العديدَ من البيتابايتات petabytes من الخزن الثانوي، وتعالج مختبرات الأبحاث الفردية روتينياً تيرابايتاتٍ من المعطيات التي تنتجها الأجهزة العلمية الخاصة بها. وكما هو الحال مع الحوسبة العالية الأداء، برز نظامٌ بيئي غني من العتاد والبرمجيات لأجل تحليلات المعطيات الكبيرة. وعلى غرار عناقيد الحوسبة العلمية، تعتمد عناقيد تحليلات المعطيات عادةً على شبكات إيثرنت السليعية والخزن المحلي، مع كون التكلفة والقدرة هما معيارا الاستمثال الأوليين. ومع ذلك، تتجه الصناعة الآن نحو شرائح FPGAs وتصاميم الشبكة المحسنة لاستمثال الأداء.

وفوق هذا العتاد، ينجز نظام أباتشي-هادوب Apache Hadoop [25] نموذج MapReduce لأجل تحليلات المعطيات. يضم Hadoop نظامَ ملفات موزع (HDFS) لإدارة عدد ضخم من الملفات الكبيرة، الموزعة (مع نسخ مكررة من الكتل) على الخزن المحلي للعنقود. إن HDFS و HBase، وهما تنجيزان مفتوحا المصدر لمخزن مفاتيح وقيم key-value

² PAPI: مختصر Performance Application Programming Interface (واجهة برمجية لأداء التطبيقات)، هي واجهة حمل portable (مكتبية الهيئة) لعدادات أداء عتادية على المعالجات الصغيرة المعاصرة. واجهة PAPI مستعملة بكثرة لانتقاط قياسات أداء منخفضة المستوى (مثل: عدد التعليمات، دورات الميقاتية، عدد مرات عدم إصابة الخابية) على منظومات حاسوبية تعمل بنظام تشغيل يونكس/لينكس. (المترجم)

يسمى BigTable من غوغل، [3] هما نظيرا المعطيات الكبيرة من Lustre للعلوم الحاسوبية، وإن كانت مستمثلة لعتاد ونماذج نفاذ مختلفة.

على قمة نظام الخزن Hadoop، أدوات (مثل Pig [18]) توفر نموذج برمجة عالي المستوى لنموذج MapReduce ذي الطورين. فإلى جانب دعم تدفق المعطيات (Storm و Flume)، والرسم البياني (Giraph)، والمعطيات العلائقية (Sqoop)، جرى تصميم النظام البيئي لـ Hadoop لتحليل المعطيات. وإضافة إلى ذلك، تتيح أدوات (مثل Mahout) إجراء عمليات تصنيف، وتوصية، وتنبؤ عن طريق التعليم بإشراف وبلا إشراف. وخلافاً للحوسبة العلمية، غالباً ما يعتمد تطوير التطبيقات لأجل تحليلات المعطيات على لغة جافا وأدوات خدمات الويب (مثل Ruby على Rails).

تصعد التحديات

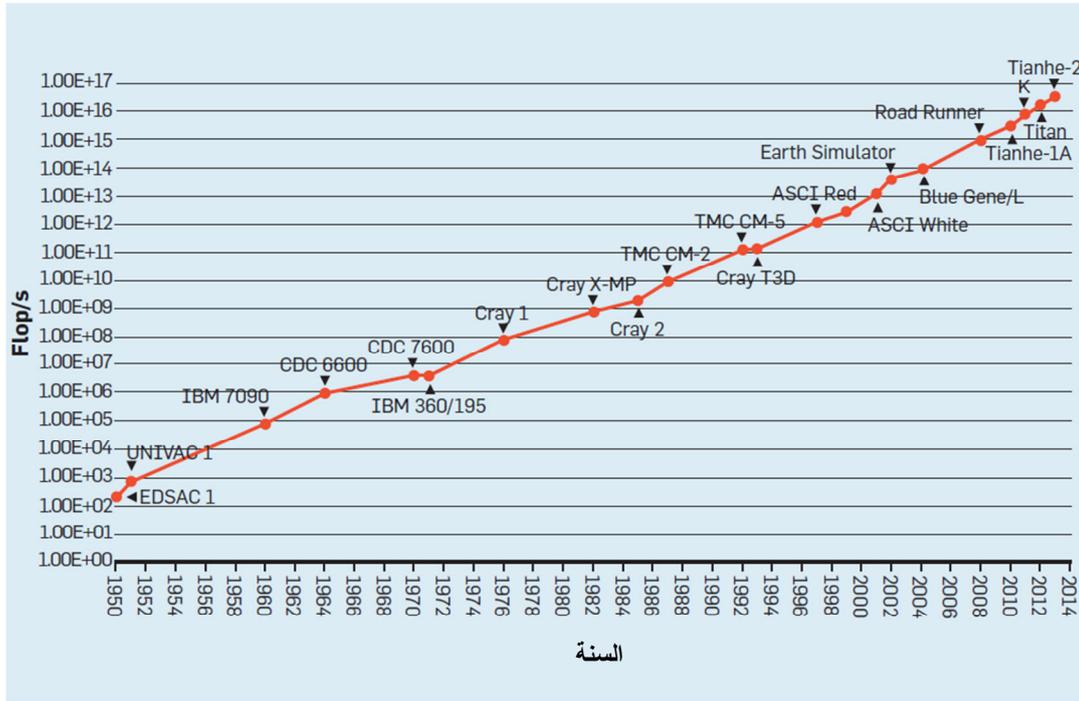
بسبب سرعة وتيرة التغير التكنولوجي (التقني)، فإن القدرة الرائدة هي هدف متحرك. الهاتف الذكي اليوم يحسب بسرعة حاسوب الأمتس الفائق، والمجموعات الموسيقية الشخصية اليوم هو كبير كبر الخزن على نطاق مؤسسة الأمتس.

ولئلا يبدو هذا من قبيل المبالغة، فإن الأداء المقيس للآي فون 6 من أبل (Apple iPhone 6) أو غالاكسي S5 من سامسونج (Samsung Galaxy S5) باستعمال برامج قياس أداء benchmark الجبر الخطي القياسية الآن يتجاوز إلى حد بعيد أداء كراي 1 (Cray-1)، الذي كان يُعدّ على نطاق واسع أول حاسوب فائق ناجح. وفي الهاتف الذكي نفسه قدرة خزن تُنافس المحتوى النصي لمكتبة بحثية رئيسية.

قبل بضع سنوات، اعتُبر التيرافلويس (10¹² teraflops عملية فاصلة عائمة في الثانية) والتيرابايت terabytes (10¹² بايت من الخزن الثانوي) أحدث ما توصلت إليه الحوسبة المتقدمة. أما اليوم، فهاتان القيمتان نفساهما تمثلان حاسوباً مكتئباً مزوداً بمسرّع إنفيديا Nvidia أو زيون فاي Xeon Phi من إنتل وخزنٍ محلي. الحوسبة المتقدمة اليوم معرّفة بنظم الحوسبة الفائقة ذات عدة بتافلويس (10¹⁵ petaflops عملية فاصلة عائمة في الثانية) ومراكز المعطيات السحابية ذات العديد من البيتابايت من الخزن الثانوي. ويبين الشكل 2 هذه الزيادة الهائلة في قدرة الحوسبة المتقدمة، استناداً إلى برنامج قياس الأداء benchmark ذي الأداء العالي (High-Performance LINPACK (HPL) [6] المستعمل على نطاق واسع وإلى القائمة Top500 لأسرع الحواسيب في العالم [16]. ومع أن حل النظم الخطية الكثيفة dense من المعادلات لم يعد القياس الأفضل للأداء الناتج على التطبيقات العلمية والهندسية المعقدة، فإن هذه المعطيات التاريخية توضح مدى سرعة تطور الحوسبة العالية الأداء. ومع أن الحوسبة العالية الأداء استفادت من التقدم في أشباه النواقل نفسه كما استفادت الحوسبة السلبية، فإن تحسن أداء النظام مستمر بسرعة أكبر نتيجة لزيادة حجم النظام والتوازي.

وقد كان نمو المعطيات الشخصية، ومعطيات الأعمال التجارية، والحكومية، والعلمية عظيمًا أكثر وموثقًا جيداً. يبني مزودو السحابة التجارية شبكات عالمية من مراكز المعطيات، كل منها تكلف مئات الملايين من الدولارات لدعم محركات البحث على شبكة الإنترنت، والشبكات الاجتماعية، والخدمات السحابية. وفي الوقت نفسه، يتحدى حجم المعطيات العلمية المنتج سنوياً الآن ميزانيات وكالات البحث الوطنية.

على سبيل المثال، يبين الشكل 3 النمو المتسارع في عدد الأغراض المخزنة في خدمة الخزن البسيط S3 من أمازون. على قمة هذه الخدمات ذات المستوى المنخفض، تقوم شركات (مثل نتفليكس Netflix) بتتجيز نظم توصية متقدمة لاقتراح أفلام على المشتركين، ثم ترسل دفعات الأفلام المختارة. والباحثون العلميون أيضاً يستكشفون على نحو

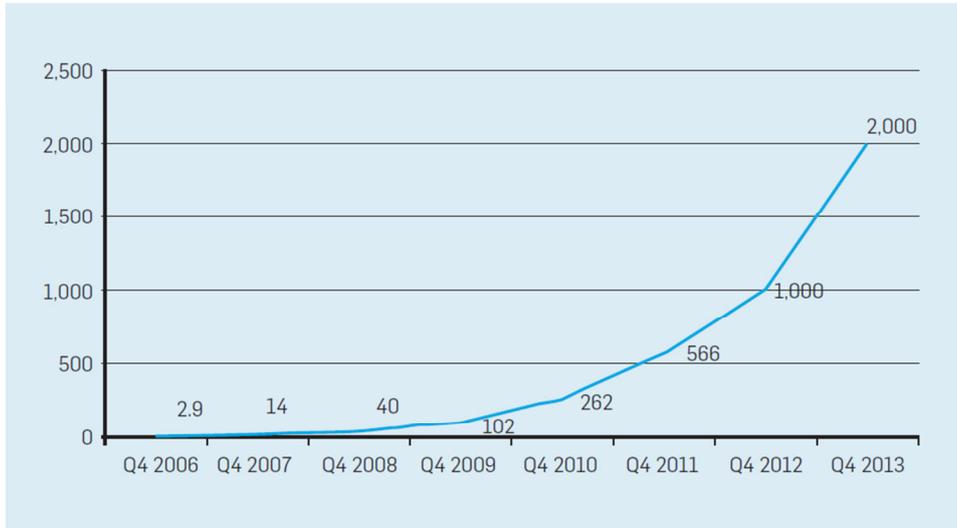


الشكل 2. أداء الحوسبة المتقدمة مقيسًا بواسطة برامج قياس الأداء HPL.

متزايد هذه الخدمات السحابية نفسها وتقنيات التعلم الآلي لاستخراج فكرة من الصور العلمية، الرسوم البيانية والمعطيات النصية. هناك تآزر تقني واقتصادي طبيعي بين التحديات التي تواجه العلوم الكثيفة المعطيات والحوسبة بمقياس إكسا، ويُعدّ التقدم في كل من الجانبين ضروريًا لتحقيق اختراقات علمية في المستقبل. يعتمد العلم الكثيف المعطيات -data-intensive science على جمع كميات هائلة من المعطيات وتحليلها وإدارتها، سواء حصلنا عليها من المحاكاة العلمية أو المرافق التجريبية. وفي كل من الحالتين، فإن الاستثمار الوطني والدولي في النظم ذات "المقياس المفرط" extreme scale سيكون ضروريًا لتحليل كميات هائلة من المعطيات التي هي الآن شائعة في العلوم والهندسة.

السباق نحو المستقبل. فيما يخص الحوسبة العلمية والهندسية، يُعدّ مقياس إكسا (10^{18} عملية فاصلة عائمة في الثانية) الوكيل المقبل في مسار طويل من زيادات الأداء الأسية التي استمرت أكثر من نصف قرن. وبالمثل، فإنّ حفظ واستدامة المعطيات على نطاق واسع داخل التخصصات وبينها، وتوليد المعطيات المترفعة والانصهار المتعدد التخصصات، والخصوصية الرقمية والأمن، كل ذلك يضع حدودًا للمعطيات الكبيرة. هذا التعريف المتعدد الأوجه للحوسبة المتقدمة يشمل أكثر من مجرد قياسات كمية لمعدلات استمرار العمليات الحسابية المستدامة أو معدلات سعة الخزن وتحليلها؛ بل هو أيضًا مصطلح نسبي يشمل التحسينات النوعية في القدرات القابلة للاستعمال لنظم الحوسبة المتقدمة على جميع المقاييس. وعلى هذا النحو، فإنه يهدف إلى اقتراح جبهة جديدة من المقدرات العملية المسلمة للباحثين العلميين والهندسيين في جميع التخصصات.

بيد أن ثمة العديد من التحديات على طريق الحوسبة الأكثر تقدمًا من أي وقت مضى، ومن ضمن ذلك، لكن لا على سبيل الحصر، استهلاك طاقة النظام، والتبريد الصديق للبيئة، والتوازي الغزير massive parallelism، وإخفاق



الشكل 3. نمو الأغراض S3 من أمازون.

المكونات، واتساق المعطيات والمبادلات transaction consistency، وإدارة المعطيات المترفعة والأنطولوجيا، والدقة (precision)، والتذكر على مقياس كبير (recall at scale)، ودمج المعطيات المتعددة التخصصات والحفاظ عليها. وقبل كل شيء، يجب ألا تصبح نظم الحوسبة المتقدمة غامضة ومعقدة لدرجة أن تكون هي وخدماتها غير صالحة للاستعمال للجميع إلا لحنفة من الخبراء. فالأدوات المفتوحة المصدر (مثل Hadoop, Mahout, Giraph)، جنباً إلى جنب مع مجموعة متمامية من الأدوات واللغات المخصصة لنطاق محدد، أتاحت التعلم الآلي للمعطيات العلمية على نطاق واسع دون معرفة عميقة لخوارزميات التعلم الآلي. وينطبق الشيء نفسه على رموز المجتمع في نمذجة العلوم الحسابية.

العتاد والبرمجيات والمعطيات والسياسات. تاريخياً، كان التقدم في الحوسبة العالية الأداء معتمداً إلى حد بعيد على التقدم المتزامن في الخوارزميات، والبرمجيات، والبنيان، والعتاد الذي أتاح مستويات أعلى من أداء الفاصلة العائمة للنماذج الحسابية. التطورات اليوم تحدها أيضاً القوات التواردية لتحليل المعطيات، وبيانات المعطيات، وأدوات التعلم الآلي التي تدير كميات ضخمة من المعطيات العلمية والهندسية.

ومع ذلك، فكما أن التغييرات في مقياس الأجهزة العلمية تجلب فرصاً جديدة، فإنها تجلب أيضاً تحديات جديدة، بعضها تقني ولكن البعض الآخر تنظيمي وثقافي واقتصادي، وليست متماثلة ذاتياً بين المقياس. واليوم، لا يمكن إنتاج نظم الحوسبة بمقياس إكسا بأسعار موثوقية معقولة، ولا يمكن لها أن تكون خاضعة لقيود هندسية واقعية على رأس المال وتكاليف التشغيل وسهولة الاستعمال والموثوقية. ولما كانت تكاليف نظم الحوسبة المتقدمة ونظم تحليل المعطيات، سواء كانت تجارية أم علمية، انتقلت من الملايين إلى المليارات من الدولارات، فإن إجراءات التصميم والقرار أصبحت بالضرورة أكثر تعقيداً ومحفوفة بالجدل. وهذا الدرس مألوف لدى أولئك في فيزياء الطاقة العالية وعلم الفلك، حيث أصبحت مسرعات الجسيمات والمقارِب (التلسكوبات) أجهزة على نطاق الكوكب (الأرضي) ومحافظات الاتحادات الدولية والسياسات العالمية. والحوسبة المتقدمة ليست استثناء.

وتوقع خبراءٌ عديدون أن تتجاوز تكاليفُ البحوث والتطوير لإنشاء نظام حوسبة بمقياس إكسا المليار دولار أمريكي، مع تكاليف تشغيل سنوية في حدود عشرات الملايين من الدولارات. وفي الوقت نفسه، هناك اعتراف متزايد بأن الحكومات ووكالات الأبحاث لم تستثمر على نحوٍ كافٍ حقاً في الاحتفاظ بالمعطيات والإدارة، ويُستدلّ على ذلك باستثمارات القطاع الخاص بمليارات الدولارات في المعطيات الكبيرة والحوسبة السحابية. تقوم أكبر مراكز المعطيات السحابية التجارية والتي تكلف كل منها أكثر من 500 مليون دولار، وجوجل، وأمازون، وميكروسوفت، وفيسبوك، وغيرها من الشركات، بتشغيل شبكاتٍ عالمية من هذه المراكز.

وفي ضوء هذه الخلفية، فإنّ دعمَ الولايات المتحدة للبحث الأساسي هو في انخفاضٍ عقدي (decade)، عندما يؤخذ التضخم النقدي في الاعتبار، [2] ولا تزال كل من الولايات المتحدة والاتحاد الأوروبي تعانين من التعافي الضعيف من الركود الاقتصادي الذي حصل عام 2008. وما يفاقم التحديات أكثر، أن السباق العالمي للهيمنة في مجال الحوسبة المتقدمة يُلْقِه رغبات الأمن القومي، والقدرة التنافسية الاقتصادية، ومستقبل النظام البيئي السائد الخاص بالحوسبة. إن التحول من الحواسيب الشخصية إلى التجهيزات قد زاد أيضاً المنافسة بين الولايات المتحدة، التي هيمنت على النظام البيئي للبنيان x86، والنظام البيئي ARM المرخص عالمياً. وفي الوقت نفسه، سببت مخاوف بشأن السيادة الوطنية، وأمن المعطيات، وحوكمة الإنترنت، منافسةً ومخاوف سياسية جديدة حول خدمات المعطيات وعمليات الحوسبة السحابية.

وبصرف النظر عن هذه التحديات، فإن هناك سبباً للتفاؤل الحذر. فقد كان كل تقدم في تقانة الحوسبة قائداً للابتكار الصناعي والنمو الاقتصادي، وهذا يشمل طيفاً كاملاً من الحوسبة، من "إنترنت الأشياء" الناشئة، مروراً بالتجهيزات النقالة الموجودة في كل مكان، إلى أقوى نظم الحوسبة في العالم وأكبر أرشيفات المعطيات. وهذه التطورات أيضاً كانت محفّزاً للبحوث الأساسية والتطبيقية في كل مجال من العلوم.

إن حل التحديات التقنية والسياسية والاقتصادية التي لا تعد ولا تحصى لن يكون سهلاً ولا حتى ممكناً بمعالجتها في عزلة. وبدلاً من ذلك، فإنه سوف يتطلب التخطيط المنسق في الحكومة، والصناعة، والأوساط الأكاديمية، والتزام التشارك في المعطيات واستدامتها، والبحث والتطوير التعاونيين، والاعتراف بأن كل من المنافسة والتعاون سيكون ضرورياً لتحقيق النجاح. يجب ألا يكون مستقبل المعطيات الكبيرة وتحليلاتها معادياً للحوسبة بمقياس إكسا؛ فكلهما بالغ الأهمية لمستقبل الحوسبة المتقدمة والاكتشافات العلمية.

الفرص العلمية والهندسية

كان الباحثون في العلوم الفيزيائية والهندسية طويلاً المستعملين الرئيسيين للحوسبة المتقدمة والنماذج الحسابية. وقد كان الدافع جزئياً وراء اعتمادها حديثاً جداً من قبل العلوم البيولوجية والبيئية والاجتماعية، صعوداً تحليلات المعطيات الكبيرة. وإضافة إلى ذلك، تُستعمل الحوسبة المتقدمة الآن على نطاق واسع في الهندسة والتصنيع المتقدم. فمن فهم دقائق تدفق الهواء في الآلات التوربينية، مروراً بديناميك الجزيئات الكيميائية للمنتجات الاستهلاكية، إلى نمذجة المواد الخام الكتلية الحيوية لخلايا الوقود، أصبحت الحوسبة المتقدمة مرادفاً للتصميم المتعدد التخصصات، والتحسين، والتصنيع المتقدم. وبالتالي، نرى حاجة إلى بعض الأمثلة فقط لتوضيح الفوائد العلمية والهندسية العميقة والمتنوعة من الحوسبة المتقدمة:

علم الأحياء (البيولوجيا) والطب الحيوي. تغيّرت طبيعة البيولوجيا والطب الحيوي (biomedicine) بالحصول على كميات ضخمة من المعطيات الوراثية (الجينية). فقد مكّنت المُسلسلات الوراثية genetic sequencers الرخيصة والعالية الإنتاجية من التقاط سلاسل الحمض النووي DNA للكائنات، ومن القيام بدراسات حول العلاقة على نطاق المجين بأمراض البشرية والتحقيقات الأحيائية الدقيقة البشرية human microbiome، وكذلك الدراسات البيئية في المجينات الفائقة metagenomics. وبوجهٍ أعمّ، تشمل التحديات البيولوجية والطبية الحيوية: حواشي التسلسل sequence ومقارنته، والتنبؤ ببنية البروتين؛ ومحاكاة الجزيئات، وآلات البروتين؛ والمسارات الاستقلابية metabolic pathways والشبكات المنظمة؛ ونماذج الخلايا الكاملة والأعضاء؛ والكائنات الحية، والبيئات، والإيكولوجيات؛

فيزياء الطاقة العالية. تُعدّ فيزياء الطاقة العالية كثيفة الحساب والمعطيات على حدّ سواء. يوفّر أول مبادئ النماذج الحسابية للصبغية الكمية quantum chromodynamics تقديراتٍ عديدة وإقراراتٍ صلاحيةٍ للنموذج القياسي". وبالمثل، تتطلب الكاشفات عن الجسيمات قياس احتمالات الأحداث "المثيرة للاهتمام" في عدد كبير من الملاحظات (مثلاً في 10^{16} أو أكثر تصادمًا للجسيمات التي تلاحظ في السنة). استلزم مصادم الجسيمات الضخم Large Hadron Collider (LHC) وتجاربه إنشاء شبكة grid حوسبة في أنحاء العالم لتبادل المعطيات واختزالها، وقيادة عملية نشر الشبكات والبروتوكولات المتقدمة، إضافةً إلى ترانزيّة من خازنات المعطيات data repositories. كلّ ذلك كان ضروريًا لتعرّف بوزون هيغز Higgs boson الذي طال انتظاره؛

علم المناخ. علم المناخ أيضًا يعتمد اعتمادًا أساسيًا على توفر بنية أساسية موثوقة لإدارة كميات كبيرة من المعطيات غير المتجانسة على نطاق عالمي والنفوذ إليها. هذا العلم بطبيعته هو جهد تعاوني ومتعدد التخصصات، ويتطلب وضع نماذج متطورة من العمليات الفيزيائية وآليات تبادل بين العوالم الأرضية، كالغلاف الجوي واليابسة والمحيطات والجليد البحري، ومقارنة وإقرار صلاحية لتلك المحاكاة باستعمال معطيات الرصد من مصادر مختلفة، وكل ذلك مجموعًا طوال مدة طويلة. ولتشجيع الاستكشاف، جعلت ناسا NASA معطيات سائل المناخ وعلوم الأرض متاحة عن طريق خدمات وب من أمازون AWS؛

علم الكونيات والفيزياء الفلكية. يعتمد علم الكونيات cosmology والفيزياء الفلكية astrophysics حاليًا اعتمادًا جوهريًا على نماذج حسابية متقدمة لفهم البنية النجمية، وتكوّن الكواكب، وتطور المجرة، والتفاعلات الأخرى. تجمع هذه النماذج بين عمليات السوائل، ونقل الإشعاع، والجاذبية النيوتونية، والفيزياء النووية، والنسبية العامة (من بين عمليات أخرى). ويكمن وراء هذه العمليات مجموعة غنية من التقنيات الحسابية المعتمدة على تنعيم شبكيّ تكيفي adaptive mesh refinement وخوارزميات جسيمات في خلية particle-in-cell³ متعددة الأقطاب multipole، وطرائق مونت كارلو، وديناميك الموائع للجسيمات المُمسّسة smoothed-particle hydrodynamics؛

علم الفلك Astronomy. واستكمالًا للحساب، توفر مسوح كامل السماء، وجيل جديد من المقاريب (التلسكوبات) المؤتمنة رؤى جديدة. وبدلًا من النقاط المعطيات الرصدية للإجابة عن سؤال معروف، يقوم علماء الفلك astronomers كثيرًا اليوم باستعلام مجموعات معطيات موجودة لاكتشاف أنماط وتوجهات لم تكن معروفة سابقًا. يُعدّ اختزال المعطيات الكبيرة ونظم التعلم بلاإشراف جزئين أساسيين من تحليل صور الاستكشاف هذا؛

³ طريقة particle-in-cell (PIC) تشير إلى تقنية مستعملة لحلّ صنف محدد من المعادلات التفاضلية الجزئية. اشتهرت هذه الطريقة في محاكاة البلازما في أواخر الخمسينيات وأوائل الستينيات من القرن الماضي. (المترجم)

علاج السرطان. يعتمد علاج السرطان الفعال على الكشف المبكر والعلاجات المستهدفة عن طريق الجراحة والإشعاع والعلاج الكيميائي chemotherapy. في المقابل، يعتمد تحديد الأورام tumor identification وتخطيط العلاج على تحسين الصورة، واستخراج السمات وتصنيفها، وتقطيعها، وتسجيلها، وإعادة بنائها بثلاثة أبعاد 3D، وتكميتها quantification. وهذه التقنيات وتقنيات التعلم الآلي الأخرى لا توفر إقرار صلاحية التشخيص فحسب، بل تُستعمل على نحوٍ متزايد لإجراء تحليلٍ مقارن وطُوليٍّ لنظم المعالجة؛

علوم المواد التجريبية والحسابية. علوم المواد التجريبية والحسابية هي المفتاح لفهم خصائص المواد والخيارات الهندسية؛ على سبيل المثال، يسمح تبعثر النيوترونات neutron scattering للباحثين بفهم بنية المواد وخصائصها، والنظم الجزيئية الكبيرة (الماكروية) macromolecular systems والبيولوجية، والفيزياء الأساسية للنترون، وذلك بتوفير معطيات عن البنية الداخلية للمواد بدءاً بالمقياس الذري atomic scale (مواضع ذرية وإثارات) وانتهاءً بالمقياس المتوسط mesoscale (مثل آثار الإجهاد)؛

إنتاج الفولاذ. يُعد إنتاج الفولاذ عن طريق الصب المستمر جزءاً مهماً من استهلاك الطاقة العالمية وإنتاج غازات الدفيئة. فحتى التحسينات الصغيرة لهذه العملية لها فوائد اجتماعية عميقة وتقصد مئات الملايين من الدولارات. تُستعمل الحواسيب العالية الأداء لتحسين فهم هذه العملية المعقدة عن طريق نماذج حسابية شاملة، ولتطبيق هذه النماذج بحثاً عن ظروف التشغيل لتحسين العملية؛

التنقيب في النصوص والمعطيات. أدى النمو الهائل من المنشورات البحثية إلى جعل العثور على البحوث ذات الصلة وتتبعها صعباً على نحوٍ متزايد. فبعيداً عن حجم النص، للمبادئ أسماء مختلفة أو متشابهة في المجالات المختلفة. ويُستعمل على نحوٍ متزايد تصنيف النصوص، وأدوات إظهار بيانية للدلالة، ونظم توصية لتعرف الموضوعات ذات الصلة واقتراح مقالات ذات الصلة بالدراسة.

هناك اثنان من المواضيع العامة المشتركة بين هذه التحديات العلمية والهندسية. الموضوع الأول هو مجموعة واسعة جداً من التفاعلات، على مقاييس زمانية ومكانية، المعقدة وغير الخطية التي تشمل عمليات بيولوجية وفيزيائية متعددة. وهذه هي الأكثر طلباً للمحاكاة الحاسوبية، وتستدعي فرقاً بحثية تعاونية، جنباً إلى جنب مع نظم حوسبة ضخمة جداً وقادرة جداً. في كل حالة، يتمثل الهدف بمحاكاة تنبؤية، أو النقاط فكرة تختبر النظريات، وتعرف تفاعلات دقيقة، وتوجه الأبحاث الجديدة.

الموضوع الثاني هو النطاق الهائل للمعطيات العلمية وتنوعها والفرص غير المسبوقة لاستيعاب المعطيات، والارتباط المتعدد التخصصات، والتحليل الإحصائي. وسواء كنا نتحدث عن العلوم البيولوجية أو الفيزيائية أو الهندسة أو الأعمال التجارية، فإن المعطيات الكبيرة تولد احتياجاتاً وفرصاً بحثية جديدة.

التحديات التقنية في الحوسبة المتقدمة

إن الفرص العلمية والهندسية التي أصبحت ممكنة بواسطة الحوسبة المتقدمة وتحليلات المعطيات هي فرص عميقة، لكن التحديات التقنية في تصميم وبناء وتشغيل نظم الحوسبة وتحليل المعطيات المتقدمة على مقياس لم يسبق له مثيل هي شاقة بالقدر ذاته. ومع أن مراكز الحوسبة السحابية والمنصات الحسابية بمقياس إكسا على ما يبدو مختلفة تماماً، كما بينا

في وقت سابق، فإن التحديات التقنية الكامنة وراء المقياس متشابهة، والعديد من الشركات والباحثين أنفسهم يستكشفون تقانات ذات استعمال مزدوج تنطبق على الموضوعين على حدّ سواء.

في سلسلة من الدراسات على مدى السنوات الخمس الماضية، تعرّفت وزارة الطاقة الأميركية عشر (10) تحديات بحثية [10, 15, 24] في استنباط جيلٍ جديد من نظم الحوسبة المتقدمة، ويتضمن ذلك ما يلي، بإضافة مقارناتنا بالحوسبة السحابية:

التقانات المُجدبة طاقياً فيما يخصّ الدارات، والقدرة الكهربائية، والتبريد. مع تقانة أشباه النواقل الحالية، تستهلك جميع التصاميم بمقياس إكسا المقترحة مئات الميجاوات من القدرة الكهربائية. وهناك حاجة إلى تصاميم وتقانات جديدة لتخفيض هذه المتطلبات من الطاقة إلى مستوى أكثر قابلية للإدارة ومُجدياً من الناحية الاقتصادية (مثل 20MW-40MW) أي مماثل لتلك المستعملة في مراكز المعطيات السحابية التجارية؛

تقانات ربطٍ عالية الأداء. في نظام الحوسبة بمقياس إكسا، تتجاوز تكلفة الطاقة لتحريك عنصر معطيات تكلفة عملية بالفاصلة العائمة، وهذا يستلزم ربطاً interconnect مُجدد طاقياً جداً ذا تلبُّثٍ منخفضٍ وعرضٍ حزمةٍ عالية لتبادل المعطيات الدقيق الحبيبية fine-grain data exchanges بين مئات الآلاف من المعالجات. وحتى مع هذه التصاميم، سوف تكون هناك حاجة إلى خوارزميات وبرمجيات واعية للمحلية (locality-aware) لتعظيم أداء الحساب والحد من الحاجة من الطاقة؛

وبدافع من ضرورات التكاليف، جرى بناء نظم الحوسبة السحابية التجارية باستعمال وصلات إترنت السلعية واعتماد نموذج حسابٍ متوازٍ متزامنٍ كثيف. ومع أن هذا النهج أثبت فعاليته، كما يتضح من الاعتماد الواسع النطاق لطقم الأدوات MapReduce (مثل Hadoop)، فإن الربطَ للتقارب المنخفض التكلفة، سيكون مفيداً لمنصات الحساب والمعطيات الكثيفة، ويتيح إمكاناتٍ جديدةً لتحليل المعطيات الدقيقة الحبيبية fine-grain data analysis.

تقانات الذاكرة المتقدمة لتحسين القدرات. إن التقليل من حركة المعطيات وتقليل استعمال الطاقة هما أيضاً يعتمدان على تقانات الذاكرة الجديدة، ومن ضمن ذلك منهج المعالج داخل الذاكرة⁴، ومنهج الذاكرة المكثفة stacked (مكعب الذاكرة HMC من شركة مَيكرون هو مثال مبكر)، ومنهج الذاكرة اللامتلاشية (nonvolatile). ومع أن التفاصيل تختلف بين الحساب وتحليل المعطيات، فإن المحددات الخوارزمية لسعة الذاكرة ستكون موجّهةً هاماً للتكلفة الإجمالية للنظام، ذلك أن الذاكرة لكل نواة core للنظم الضخمة جداً سوف تكون بالضرورة أصغر مما هي عليه في التصاميم الحالية؛

برمجيات النظام القابلة للتصدّد الواعية للطاقة والأعطال. استندت برمجيات الحوسبة العالية الأداء التقليدية إلى افتراض أن الأعطال نادرة؛ ومع اقترابنا من مستويات مقياس إكسا، فإن صمود النظام في مواجهة الأعطال الاعتيادية للمكونات سوف يكون أمراً أساسياً. وبالمثل، يجب أن تصبح إدارة الطاقة الديناميكية والتكيفية جزءاً لا يتجزأ من برمجيات النظام، لأسباب اقتصادية وفنية على حدّ سواء.

⁴ المعالج داخل الذاكرة (processor in memory)، ويُسمى أحياناً المعالجة داخل الذاكرة (processing in memory) أو اختصاراً (PIM)، هو إدماج المعالج في ذاكرة الرام RAM على شريحة chip واحدة. ويطلق على النتيجة أحياناً رقاقة PIM. في هذه الشريحة يكون منطق المعالج مربوطاً مباشرة بالذاكرة، وهذا يزيد من سرعة المعالجة ومعدل النقل إلى الذاكرة ويقلل من التلبُّث (latency) واستعمال الطاقة، ومن ثم يتغلب على عنق الزجاجة في نموذج فون نويمان حيث المعالج والذاكرة منفصلان والمعطيات تنتقل بينهما مع وجود تلبُّث. (المترجم)

إنّ الخدمات السحابية لتحليلات المعطيات، تبعاً لاتفاقات جودة الخدمة التجارية الخاصة بها، تُجسّد أعداداً كبيرة من تقنيات المرونة، ومن ذلك التوزيع الجغرافي، وإعادة التشغيل وتجاوز الأعطال failover آلياً، وحقن الأعطال failure injection، والمراقبة التأملية introspective monitoring؛ ويُعدّ "Simian Army" من نتفليكس Netflix⁵ أنموذجاً توضيحياً لهذه التقنيات.

برمجيات إدارة المعطيات التي يمكنها التعامل مع الحجم، والسرعة، وتنوع المعطيات. سواء وُلدت المعطيات حسابياً أو النُقِطت من الأجهزة العلمية، فإن تحليل تلك المعطيات بكفاءة في الموقع يتطلب إعادة هيكلة تدفقات العمل العلمية والتطبيقات، بالبناء على الدروس المستفادة من قنوات توارد pipelines تحليل المعطيات التجارية، فضلاً عن تقنيات جديدة لتنسيق المعطيات، وتعلّمها، والتنقيب فيها. وبدونها، سوف تُحدّ اختناقات الدخل/الخرج من نفع النظام وإمكان تطبيقه؛

برمجة نماذج للتعبير عن التوازي الغزير، ومحلية البيانات والمرونة. يضع نموذجُ الإجراءات المتتابعة ذات التواصل communicating sequential process المستعملُ على نطاق واسع، أو البرمجة باستعمال MPI، عبء المحلّة والموازاة على مطوري التطبيقات. وستشهد نظم الحوسبة بمقياس إكسا توازياً بمليار طريقة وأخطاءً متكررة. وما نحتاج إليه هو نماذج برمجة أكثر تعبيراً قادرة على التعامل مع هذا السلوك وتبسيط جهود المطور مع دعم التوازي الدقيق الحبيبية fine-grain parallelism الديناميكي.

يمكن تعلم الكثير من خدمات الوب والسحابة حيث طبقات التجريد وأطقم الأدوات المتخصصة تسمح للمطورين بنشر بيانات تنفيذ حسب الطلب (آلات افتراضية) ورفع خدمات رفيعة المستوى للحد من المعطيات المعقدة. يُبقي تحدي الحوسبة العلمية على التعبيرية والإنتاجية وفي الوقت ذاته أيضاً يوفر الأداء العالي.

إعادة صياغة المشاكل العلمية وإعادة بناء خوارزميات الحل. جرى استثمار عدة آلاف شخص-سنة في الأرمزة (codes) العلمية والهندسية الحالية وفي التنقيب في المعطيات والبرمجيات التعليمية. سوف يتطلب تكييف الأرمزة العلمية للتوازي بمليارات الطرق إعادة تصميم الخوارزميات، أو حتى إعادة ابتكارها، وربما أيضاً إعادة صياغة المسائل العلمية. إنّ مكاملة برمجيات تحليلات المعطيات والأدوات بالحساب هو أمر شاقّ أيضاً؛ فغات البرمجة ونماذجها تختلف، كما هو حال المجتمعات والثقافات. إنّ فهم كيفية القيام بهذه الأشياء بكفاءة وفعالية سيكون مفتاح حل مشاكل العلم ذات المهام الحرجة.

ضمان الصحة في مواجهة الأعطال، وقابلية إعادة الإنتاج، والتحقق الخوارزمي. مع الأعطال العابرة والدائمة المتكررة، ونقص قابلية إعادة الإنتاج reproducibility في الاتصال الجماعي، والخوارزميات الرياضية الجديدة ذات التحقق المحدود، فإنّ إقرار صلاحية الحساب وضمان الصحة سوف يكون أكثر أهمية بكثير للجيل القادم من نظم التوازي الغزير، سواء كانت مستمثلة للحوسبة العلمية، أو تحليل المعطيات، أو لكليهما؛

الاستمثال الرياضي وتكمية عدم اليقين (الشك) في حالة الاكتشاف، والتصميم، والقرار. إنّ الحسابات على نطاق واسع هي نفسها التجارب التي تُسبر فضاء العينات من النماذج العددية. إنّ فهم حساسية التوقعات الحسابية لنموذج المدخلات والافتراضات، وخاصة عندما تتطوي على تطبيقات معقدة متعددة التخصصات، يتطلب أدوات وتقنيات جديدة

⁵ <http://techblog.netflix.com/2011/07/netflixsimian-army.html>

لإقرار صلاحية التطبيق وتقييمه. النظران الهامان بالقدر نفسه في تحليلات المعطيات على نطاق واسع والتعلم الآلي هما الدقة precision (الجزء من المعطيات المسترجعة التي هي ذات صلة) والتذكر recall (الجزء المسترجع من المعطيات ذات الصلة)؛

هندسة البرمجيات والبنى الداعمة للتمكين للإنتاجية. مع أن أدوات البرمجة، والمترجمات (المصرفات)، والمفليات debuggers، وأدوات تحسين الأداء تُحدد إنتاجية البحوث لجميع نُظم الحوسبة، فإنه على مقياس كبير (at scale)، سيكون تصميم التطبيقات وإدارتها لتحقيق حسابٍ موثوق وفعال وصحيح أمرًا شاقًا جدًا. وما لم تزدد إنتاجية الباحثين، فإن الوقت اللازم لإيجاد حل يمكن أن يسيطر عليه تطوير التطبيقات، لا الحساب.

حددت دراسات عتادية وبرمجية مماثلة [1, 14]، أُجريت لمصلحة وكالة مشاريع البحوث المتقدمة الدفاعية في الولايات المتحدة، التحديات التالية، المماثلة كثيرًا لتلك التي ذكرتها دراسات وزارة الطاقة: تشغيل مُجدٍ طاقيًا. إنَّ التشغيل المُجدي طاقيًا energy-efficient operation لتحقيق معدلات الحساب المرغوبة خاضع للطاقة المبددة إجمالاً؛

سعة الذاكرة. إنَّ سعة الذاكرة الأولية والثانوية ومعدلات النفاذ، خاضعة لقيود الطاقة؛ التساير والمحلية. هما التساير والمحلية لبلوغ أهداف الأداء وفي الوقت نفسه السماح لبعض النياسب threads بالتوقف خلال العمليات ذات التلبث الطويل؛

المرونة Resilience. هي المرونة أخذًا بالحسبان للعدد الضخم للمكونات، والحجوم المتناقصة لطبيعة السيليكون، والتشغيل المنخفض الطاقة، والأعطال العابرة والدائمة للمكونات؛ تصعد التطبيقات. إنَّ تصعد التطبيقات application scaling خاضع لقيود سعة الذاكرة وتلبث الاتصال؛ إدارة التوازي. وهي الإعراب عن إدارة التوازي والمحلية في برمجيات النظام ونماذج البرمجة المحمولة، ويشمل ذلك نظم زمن التنفيذ runtime، والمُجدولات schedulers، والمكتبات؛ أدوات برمجية. أدوات برمجية لضبط الأداء، وتقييم الصحة، وإدارة الطاقة.

وإضافة إلى ذلك، وجدت دراسة عام 2011 قامت بها الأكاديمية الأمريكية الوطنية للعلوم (NAS) [9]، وهذا يستبعد حدوث فتح علمي، أن الزيادات الهائلة في الأداء المستمدة من تقلص حجم خاصية أشباه النواقل semiconductor والابتكار البناني قد شارفا على نهايتهما. وترى هذه الدراسة، جنباً إلى جنب مع دراسات أخرى، أن تقانة الحوسبة ينتظرها نقاط انعطاف هامة، على المقياس الأكبر نطاقاً (الحوسبة العالية الأداء الرائدة)، وعلى المقياس الأصغر نطاقاً (عمليات أشباه النواقل). ولا يزال مجتمع الحوسبة منقسمًا بشأن المناهج الممكنة، فمنهم الذين لديهم اعتقاد قوي أن العقبات التقنية التي تُحد من توسيع المناهج الحالية سوف يُغلب عليها، ومنهم الذين يعتقدون وجود حاجة إلى تقانات ومناهج تصميم أكثر تطرفاً (مثل تجهيزات الكم quantum والناقلية الفائقة superconducting).

تحديات العتاد والبنان. مع أن تقديم وصف كامل للتحديات التقنية المتعلقة بالعتاديات والبرمجيات التي وصفناها آنفاً يقع خارج نطاق هذا المقال، فإن استعراض مجموعة جزئية منتقاة يُعد مفيداً بقصد إلقاء الضوء على عمق واتساع المشاكل وانعكاساتها على مستقبل كل من الحوسبة المتقدمة والنشر الأوسع لتقانات الجيل القادم من حوسبة المستهلك وحوسبة الأعمال التجارية.

تصعد ما بعد Dennard. على مدى عقود، بقي "قانون" مور Moore's law صحيحاً بسبب العمل الشاق والإبداع من عدد كبير من الناس، وكذلك العديد من مليارات الدولارات المستثمرة في تقانة العمليات ومسابك السيليكون. واعتمد أيضاً على مبدأ تصعد دينر⁶ Dennard scaling، [5, 13]، الذي يوفر وصفاً لتقليص الترانزستورات ويعطي دارات أصغر حجماً بكثافة الطاقة نفسها. أدى تقليص الحجم الخطي للترانزستور مرتين إلى تقليص الطاقة أربع مرات، أو تقليص كل من الجهد (الفلطية) والتيار إلى النصف.

ومع أن حجم الترانزستور يستمر في الانكماش، إذ يسود اليوم الحجم النموذجي 22 نانومتر، فإن استهلاك الترانزستور للطاقة لم يعد ينقص تبعاً لذلك. وأدى ذلك إلى فرض قيود على معدل ميقانية الرقاقة واستهلاك الطاقة، جنباً إلى جنب مع تصميم الرقاقات المتعددة النوى multicore chips وصعود السيليكون المعتم dark silicon (رقاقات تحوي عدداً من الترانزستورات يفوق العدد الذي يمكن أن يكون نشطاً في وقت واحد، بسبب قيود حرارية وطاقيه. [8]) وقد حفزت تحديات أشباه النواقل هذه إلى إعادة النظر في تصميم الرقاقات، إذ يجب أن تكون ميزات الأداء المحتملة للحيل البنائية (التوارد الفائق super-pipelining، و score-boarding⁷، والتحويل إلى منتجات vectorization، والموازاة parallelization) متوازنة إزاء استهلاكها للطاقة. غالباً ما تُسفر التصميم البسيطة والمسرعّات المحددة الوظيفة عن توازن أفضل بين استهلاك الطاقة والأداء. وهذا التحول البنائي سوف يكون صحيحاً خاصةً إذا كان ميزان العمليات على الأعداد الصحيحة integer، وعمليات التفرع branching، والعمليات بالفاصلة العائمة يتحول لدعم تحليل المعطيات والحوسبة في الموقع.

وقد حفزت حدود طاقة الرقاقة إلى الاهتمام الكبير بالنظام البيئي للمعالج ARM. فبسبب من استمثال تصاميم المعالجات ARM للاستعمال في التجهيزات المضمنة والنقالة، إذ كان استهلاك الطاقة المحدود موجّه التصميم مدةً طويلة، كان لدى تلك المعالجات قنوات توارد ومفككات ترميز أبسط منها في تصاميم المعالجات x86.

في هذا العالم الجديد، يُعدّ التصميم المشترك للعتاد والبرمجيات ضرورة، حيث تترابط التجهيزات وبرمجيات النظم فيما بينها. ومن آثار ذلك: زيادات في الأداء المتعدد الأغراض أقل كثيراً، وتنوع في العتاد أكثر، وارتفاع في الاستمثال

⁶ هو قانون تصعد مفاده أن كثافة القدرة تبقى ثابتة مع زيادة انخفاض حجم الترانزستورات، بمعنى أن استهلاك القدرة يبقى متناسباً مع المساحة: كلا الفلطية والتيار يتصعدان (هيوطاً) مع الطول. يرتبط قانون تصعد دينر بـ (1) قانون مور Moore الذي يفيد بأن انخفاض حجم الترانزستور يسمح بزيادة عدد الترانزستورات أكثر فأكثر في الرقاقة chip دون زيادة في التكلفة، و (2) الادعاء أن أداء الحوسبة مقيساً بالواط يزداد أسياً بالمعدل ذاته تقريباً.

أتاح انخفاض قدرة الترانزستور للمصنّعين زيادة ترددات الميقانية كثيراً من جيل إلى الجيل التالي دون إحداث زيادة هامة في استهلاك الدارة الإجمالي للقدرة.

واعتباراً من 2005-2007 بدا تصعد دينر أنه انهار؛ فمع أن تعداد الترانزستورات في الدارات المتكاملة يزداد، فإن التحسن الناتج في الأداء منخفض ولا يتناسب مع التسريع الناتج عن الزيادة الكبيرة في التردد. والسبب الرئيسي لهذا الانهيار هو أنه عند الحجوم الصغيرة، تُسبب تسريبات التيار تحدياتٍ أعظم، وتُجعل الرقاقة تُسخن، وهذا ما يوّد تهديداً بتسرّب حراري ويزيد تكاليف الطاقة.

وقد أدى انهيار تصعد دينر، وتبعاً لذلك عدم المقدر على زيادة تردد الميقانية كثيراً، إلى تحوّل معظم مصنّعي وحدات المعالجة إلى التركيز على المعالجات المتعددة النوى multicore وسيلةً بديلةً لتحسين الأداء. (المترجم)

⁷ هي طريقة مركزية تُستعمل لجدولة قناة توارد pipeline ديناميكياً بحيث يمكن تنفيذ التعليمات من دون ترتيب out-of-order عندما لا توجد نزاعات ويكون الأمر متاحاً على مستوى العتاد. (المترجم)

المتعدد المتغيرات (مثل الطاقة، والأداء، والموثوقية) في نماذج البرمجة، وتحديات جديدة على صعيد إدارة موارد برمجيات النظام.

المرونة وفعالية الطاقة على مقياس كبير. بسبب من أنّ النظم المتقدمة للحوسبة وتحليل المعطيات تنمو على الدوام، فإن افتراض التشغيل بموثوقية تامة يصبح أقل صدقية بكثير. ومع أن الزمن الوسطي قبل وقوع عطل (MTBF) للمكونات المنفردة يتزايد باطراد تدريجيًا، فإن العدد الإجمالي الكبير للمكونات في هذه النظم هو دليل على أنّ النظم ذاتها سوف تُخفق على نحوٍ أكثر تواترًا. أثبتت التجارب حتى اليوم إمكان احتواء الأعطال، لكن بشرط توفر تقنيات محسنة لكشف وفهم أعطال المكونات.

وتدل مراكز المعطيات السحابية التجارية على أنّ بعض الافتراضات التي كانت سائدة مدة طويلة بخصوص أعطال المكونات وعمرها غير صحيحة. [11, 20-22] وأظهرت دراسة لجوجل عام 2009 [22] أنّ معدلات أخطاء ذاكرة الرام الديناميكية DRAM كانت أعلى بمراتب عديدة من الكبر مما صرّح عنه سابقًا، إذ إنّ ما يزيد على 8% من ذاكرة DIMM شملتها الأخطاء في السنة. ومن المدهش أيضًا أن هذه الأخطاء كانت فادحة hard، بدلًا من هينة، ويمكن تصحيحها (باستعمال رماز مصحح للخطأ).

وإضافة إلى المرونة، يجلب المقياس الكبير scale أيضًا تحديات جديدة في مجال إدارة الطاقة والتبديد الحراري. إنّ نظم الحوسبة المتقدمة وتحليل المعطيات اليوم تستهلك عدة ميغاواط من القدرة الكهربائية، وقدرة تبريد وذروة حمل طاقة حديثين إذ يمكن وضع عدة نظم في أماكن متباعدة جغرافيًا. وكما تعلم مشغلو السحابة التجارية، فإنّ البنية الأساسية للطاقة والكهرباء يمثلان الجزء الهام من التكلفة الإجمالية للنظام على المقياس الكبير، وهذا يستلزم مناهج بنية أساسية ونماذج تشغيل جديدة، ويتضمن ذلك تصاميم منخفضة الطاقة، ومناهج تبريد، وحساب تكاليف الطاقة، وكفاءة التشغيل.

التحديات البرمجية والخوارزمية. العديد من التحديات البرمجية والخوارزمية التي تواجه الحوسبة المتقدمة وتحليلات المعطيات الكبيرة هي نفسها عواقب المقياس الحدي للنظام. وكما لوحظ في وقت سابق، تُسهم الحوسبة العلمية المتقدمة في العديد من مشاكل التصعد التي تواجه خدمات الوب والسحاب ولكن تختلف في توازن استمثال سعرها-أدائها، مع التشديد على مستويات عالية من الأداء، سواء للحساب أو لتحليل المعطيات. وهذا التمييز أساسي لخيارات التصميم ومعايير الاستمثال. وانطلاقًا من المقياس الكبير ومعدلات الخطأ المتوقعة لنظم الحوسبة بكقياس إكسا، يجب إعادة النظر في تصميم وتجزير الخوارزميات من المبادئ الأولى، ويشمل ذلك استكشاف الخوارزميات العديمة التزامن الشمولي (أو على الأقل ذات الحد الأدنى منه)، وخوارزميات التغاضي عن الأعطال fault-oblivious والتسامح في الأخطاء-error-tolerant، والخوارزميات الواعية للبنيان المناسبة للعناد غير المتجانس وذي التنظيم التراتبي، ودعم العمليات الحسابية المختلطة الدقة، وبرمجيات الحوسبة ذات الكفاءة في استعمال الطاقة.

المحلية والمقياس الكبير. كما أشرنا سابقًا، من المتوقع أن تتطلب التصاميم المفترضة لنظم الحوسبة ذات المقياس المفرط extreme scale تسايرًا حسابيًا بمليارات الطرق، مع توازن قوي على جميع مستويات النظام. وسوف يكون الحفاظ على توازن الحمل على جميع مستويات التسلسل الهرمي للخوارزميات والمنصات مفتاح التنفيذ الفعال. وهذا من شأنه أن يتطلب آليات زمن تشغيل ديناميكية ومتكيفة [4]، وتخصيص موارد واعيًا للذات self-aware بغية تحمل ليس فقط الاختلالات الخوارزمية، بل أيضًا التغير في أداء العناد وموثوقيته.

في مقابل ذلك، سوف تفرض تكاليف الطاقة وتلبّاتُ الاتصال استثناءً أكبر لمحلية الحساب مما هو عليه اليوم. إن قلب النماذج التي صمدت مدة طويلة، والعمليات الحسابية سوف يكون مستهلكاً للطاقة أقل كثيراً من الاتصالات وأشدّ كفايةً منها. وعادة ما يعبر عن التعقيد الخوارزمي بدلالة عدد العمليات التي تجرى بدلاً من حركة المعطيات إلى الذاكرة. وهذا يتعارض مباشرة مع التكاليف المتوقعة للحساب على نطاق واسع، إذ سوف تكون حركة الذاكرة مكلفة جداً والعمليات من دون تكلفة تقريباً، وهي مسألة ذات أهمية للخوارزميات الكثيفة العمليات بالفاصلة العائمة، وتحليل المعطيات.

سوف تتحدى التكلفة الزمنية لحركة المعطيات مناهج التصميم الخوارزمي التقليدية والاستمثالات النسبية، وهذا ما يجعل الحساب الفائض في بعض الأحيان مفضلاً للتشارك في المعطيات، ورفع تعقيد الاتصالات إلى مستوى التعادل مع الحساب. ولذلك فمن المهم لجميع علوم الحاسوب أن يجري تصميم الخوارزميات التي تتصل بأقل قدر ممكن، للوصول من الناحية المثالية إلى الحدود الدنيا من مقدار الاتصالات المطلوبة. وسيتطلب ذلك أيضاً نماذج وأساليب لتقليل التلبّات والتسامح به (إخفائه)، واستمثال حركة المعطيات، وإزالة التزامن الشمولي.

برمجيات نظام قادرة على التكيف. لا تزال إدارة الموارد لنظم الحوسبة العالية الأداء الحالية متجزرة في نموذج خارق ذي جدولة منسقة واتصالات متزامنة بإحكام. ومع ذلك، فإن قيود: المقياس المفرط extreme scale، وعدم التجانس العتادي، ومقدرة النظام، والتبديد الحراري، وزيادة معدلات أعطال المكونات، كلّ ذلك يؤثر ليس فقط في تصميم التطبيقات وتجزئتها، بل في تصميم برمجيات النظام في مجالات متنوعة مثل إدارة الطاقة والدخل/الخرج. وبالمثل، مع ازدياد حجم المعطيات العلمية، فإنه من غير الواضح إمكان تصدّد تجريدات الملفات التقليدية ونظم الملفات المتوازية التي تستعملها الحوسبة التقنية لتتوافق مع تحليل المعطيات بمقياس بيتا petascale.

وبدلاً من ذلك، سوف يتطلب الأمر أن تدعم التصميمات الجديدة لبرمجيات النظام ونظام التشغيل، إدارة الموارد غير المتجانسة وتراتبيات الذاكرات نوات الخوابي غير المتسقة non-cache-coherent memory hierarchies، وأن تُهيئ للتطبيقات وبيئات وقت التشغيل runtime مزيداً من السيطرة على سياسات جدولة المهام، وأن تدير فضاءات الأسماء الشمولية. وأيضاً أن تكشف آليات لقياس أدق، وتنبؤ، وسيطرة على إدارة الطاقة أكثر دقة، كي تتمكن برامج الجدولة من إتباع الحسابات بالسرعات المحددة الوظيفية، وإدارة الأغذية الحرارية وملاحم الطاقة للتطبيقات. وقد واجه مزودو السحابة التجارية بالفعل العديد من هذه المشاكل، وخبرتهم في إدارة الموارد على نطاق واسع لديها الكثير لتقدمه للنظام البيئي الحوسبة العلمية.

دعم البرمجة المتوازية. كما زاد تنوع عتاد الحوسبة المتقدمة وتعقيده ومقياسه، فقد زاد أيضاً تعقيد تطوير التطبيقات وصعوبته، إذ العديد من وظائف التشغيل تُعَيَّن الآن للتطبيقات. ونفاقم تعقيد التطبيقات على نحو متزايد بسبب الطبيعة المتعددة التخصصات للتطبيقات التي تجمع بين الخوارزميات والنماذج التي تمتد على مجموعة من المقاييس الزمانية المكانية والمناهج الخوارزمية.

لنتأمل نموذج البرمجة المتوازية ذات البرنامج المنفرد والمعطيات المتعددة أو النموذج المتوازي الكثير التزامن، حيث تقسم معطيات التطبيق وتوزع على الذاكرات أو الأقراص المنفردة العائدة لعقد الحساب، وتتشارك العقد في المعطيات بواسطة التراسل على الشبكة. في المقابل، يدير رموز التطبيق في كل عقدة تراتبية [عتاد] الحساب المحلي المتعددة المستويات (وهي عادةً عدة نوى cores متعددة النياسب multithreaded وربما غير متجانسة، وغالباً) مسرّع GPU وحدة معالجة بيانية، وينسق عمليات الدخل/الخرج I/O، ويدير تدقيق التطبيق، ويشرف على ميزانيات الطاقة والتبديد الحراري. وهذا

المستوى الشاق من التعقيد، والتشكيل والضبط المفصل، يجعل تطوير التطبيقات المنيعة فناً غامضاً في متناول قلةٍ متمكنةٍ ومتفانيةٍ فقط.

من الناحية المثالية، سوف يرفعُ التصميمُ المستقبلي للبرمجيات، وتطويرها، ونشرها مستوى التجريد ويدخل الأداء والصحة في الحسبان من البداية بدلاً من النهاية. وإضافة إلى تصميم وتطوير التطبيقات أكثر وعياً للأداء باعتماد نماذج أداءٍ وصحةٍ متكاملة، يجب أن تكون هذه الأدوات متكاملةً مع المترجمات (المصرّفات) ونظم وقت التشغيل، وأن توفر المزيد من الدعم للعتاد غير المتجانس ونماذج البرمجة المختلطة، وتقدّم معالجةً وتحليلاً للمعطيات أكثر تطوراً.

ولعل نماذج البرمجة وأدواتها هما أكبر نقطة للتباعد بين الحوسبة العلمية والنظم البيئية للمعطيات الكبيرة. فالأخيرة تركّز الاهتمام في التجريدات البسيطة (مثل مخازن المفاتيح-القيم و MapReduce)، جنباً إلى جنب مع مصاعغات المعطيات الغنية الدلالة والمواصفات العالية المستوى. وقد سمح ذلك للعديد من المطورين بإنشاء تطبيقاتٍ تعلّم آلي معقدة مع القليل من المعرفة للعتاد الأساسي أو برمجيات النظام. في المقابل، واصلت الحوسبة العلمية الاعتماد إلى حدٍ بعيدٍ على اللغات التقليدية والمكتبات.

سوف تكون لغات البرمجة ونماذج البرمجة الجديدة، فيما عدا C و FORTRAN، عوناً (للمبرمجين). ولما كانت برمجيات التطبيقات بالفعل حاضرةً للحوسبة التقنية، فإنّ الابتعاد الجذري ليس منهجاً واقعياً. لقد كان لميزات البرمجة الموجودة في لغات جديدة (مثل شابل Chapel و X10) بالفعل تأثير غير مباشر في نماذج البرامج القائمة. وقد استفادت نماذج البرمجة الحالية (مثل OpenMP) سلفاً بالتوسيعات الأخيرة المتعلقة، مثلاً، بتوازي المهام task parallelism، والمسرّعات، والتصاق النياسب thread affinity.

اللغات المحددة النطاق (DSLs) هي لغات تختص بمجال تطبيقٍ معيّن، وتُعدّ وسيلةً لتوسيع لغة قاعدية موجودة بتضيق توسيعات اللغات DSL. إن اللغات DSLs المضمّنة هي وسيلةٌ عمليةٌ لاستغلال قدرات التحليل والتحويل المعقدة التي تتمتع بها مترجمات (مصرّفات) اللغات القياسية. إذ يكتب المطور التطبيق باستعمال بدئيات عالية المستوى high-level primitives، فيحوّله المترجم إلى رمازٍ منخفض المستوى ذي كفاءة كي يحسّن الأداء على المنصة الأساسية. **تحديات خوارزمية ورياضية.** سوف تتسبب الحوسبة بمقياس إكسا بزيادة الطلب على الخوارزميات في مجالين على الأقل: الحاجة إلى كميات متزايدة من محلية المعطيات لتنفيذ العمليات الحسابية بكفاءة، والحاجة إلى الحصول على مستوياتٍ توازي دقيقةٍ الحُببية أعلى بكثير، ذلك أنّ النظم الراقية تدعم أعداداً متزايدة من نياسب الحساب compute threads. ونتيجة لذلك، يجب أن تتكيف الخوارزميات المتوازية مع هذه البيئة، ويجب تطوير خوارزمياتٍ وتجزياتٍ جديدة لاستغلال القدرات الحسابية للعتاد الجديد.

إنّ التطوير الهام للنموذج، وإعادة التصميم الخوارزمي، وإعادة تجزير التطبيقات العلمية، الذي يدعمه نموذج (أو نماذج) البرمجة المناسبة بمقياس إكسا، كل ذلك سوف يكون ضرورياً بغية استغلال قوة البنى ذات مقياس إكسا. وسوف يكون الانتقال من الحوسبة بمقياس بيتا الجزئي وبمقياس بيتا الحاليين، إلى الحوسبة بمقياس إكسا على الأقل شاقاً بالقدر ذاته للنقلة من حوسبة المتجهات إلى الحوسبة المتوازية التي جرت في التسعينيات من القرن الماضي.

التحديات الاقتصادية والسياسية

يحدّد ملامح التحديات التقنية للحوسبة المتقدمة وتحليلات المعطيات الكبيرة عناصر أخرى من مشهد الحوسبة الأوسع. وبوجه خاص، بوجود الهواتف الذكية القوية وخدمات الحوسبة السحابية لم تُعد الحواسيب الشخصية والخدمات المحلية مقيس الحوسبة. وأثار هذا التحول أيضاً منافسةً دولية في الميزات الصناعية والتجارية، مع استثمار الدول والمناطق في تقانات جديدة وتوطينات جديدة للنظم.

تحولات النظام البيئي للحوسبة. إنّ ثورة الإنترنت وخدمات الويب هي ثورة عالمية، وإنّ تأثير الولايات المتحدة، وإن كان كبيراً، في تساؤل. وبصرف النظر عن نجاح شركة أبل Apple الهائل، فإنّ معظم الهواتف الذكية والأجهزة اللوحية تصمّم الآن وتُبنى وتُفتتى على الصعيد العالمي، وإنّ حجم المبيعات السنوية للهواتف الذكية والأجهزة اللوحية يفوق الحجم الخاص بالحواسيب الشخصية والخدمات.

ويرافق هذا التحول المستمر في أذواق المستهلكين وفي الأسواق تحولاً تقني آخر. تعتمد الهواتف الذكية والأجهزة اللوحية على معالجات صغيرة اقتصادية في الطاقة (وهذا مكون أساسي لتصاميم الحوسبة بمقياس إكسا المقترحة)، وعلى منظومات في رقاقة (SoCs) باستعمال بنين المعالج ARM. وخلافاً لشركتي إنتل و ADM، اللتين تصممان وتصنعان شرائح المعالجات x86 الموجودة في الحواسيب الشخصية الحالية ومعظم الخدمات ونظم HPC الرائدة، فإنّ شركة ARM لا تصنّع الرقاقت الخاصة بها. وبدلاً من ذلك، ترخّص تصميماتها للآخرين، الذين يدمجون بنين المعالج ARM في شريحة SoCs مخصّصة بحسب الطلب تُصنّعها مصانع سبك أشباه النواقل العالمية، مثل شركة Taiwan TSMC (Semiconductor Manufacturing Company) التايوانية.

مشاريع دولية بمقياس إكسا. إنّ السباق الدولي المحيط بالحوسبة المتقدمة يجمع بين الاهتمام بكل من القدرة التنافسية الاقتصادية، وتحول النظام البيئي التقني (مثل نظم ARM و x86)، وحوسبة الأعمال التجارية والحوسبة التقنية (مثل خدمات الحوسبة السحابية ومراكز المعطيات)، والبحوث العلمية والهندسية. أطلق الاتحاد الأوروبي واليابان والصين والولايات المتحدة جميعهم مشاريع حوسبة بمقياس إكسا، ويختلف اهتمام كل منها فيما يخص تقنيات العتاد، وبرمجيات النظام، والخوارزميات، والتطبيقات.

الاتحاد الأوروبي. أعلن الاتحاد الأوروبي (EU) عن بداية برنامجه البحثي بمقياس إكسا في تشرين الأول (أكتوبر) 2011 بتمويلٍ مقداره 25 مليون يورو يشمل ثلاثة مشروعات بحثية يتم بعضها بعضاً، في جهوده في إطار العمل السابع (Framework 7). الغرض من كل من المشاريع التالية: البحوث التعاونية نحو النظم والأدوات والتطبيقات بمقياس إكسا (CRESTA: Collaborative Research into Exascale Systemware, Tools and Applications)، ومنصة الدخول الديناميكية إلى مقياس إكسا (DEEP: Dynamical Exascale Entry Platform)، ومشاريع مون بلان (Mont-Blanc)، هو التحقيق، كل على حدة، في تحديات المقياس إكسا المختلفة باستعمال نموذج تصميم مشترك يشمل: العتاد، وبرمجيات النظام، والتطبيقات البرمجية. وتمثل هذه المبادرة الاستثمار المستدام الأول على المستوى الأوروبي في بحوث مقياس إكسا.

يجمع المشروعُ CRESTA أربعة مراكز أوروبية للحوسبة العالية الأداء: مركز إدنبره Edinburgh للحوسبة المتوازية (رائد المشروع)، مركز الحوسبة العالية الأداء-شتوتغارت Stuttgart، مركز IT للعلوم المحدود-فنلندا، ومركز التطوير الشريك-السويد، إضافة إلى جامعة دريسدن Dresden للتقانة، التي سوف تقدم الخبرة في استمثال الأداء. وإضافة إلى ذلك، يتضمن فريق CRESTA أيضاً مهنيّ التطبيقات من العلم والصناعة الأوروبية، فضلاً على باعة الحوسبة العالية

الأداء HPC، ومنهم مطور أدوات HPC، Allinea، وبائع HPC، كراي Cray. يركّز الفريق CRESTA على استعمال التطبيقات بوصفها موجهًا للتصميم المشترك لبيئات تطوير البرمجيات، والخوارزميات والمكتبات، وأدوات المستعمل، والتقانات الأساسية والبيئية.

يجمع مشروع Mont-Blanc، بقيادة مركز الحوسبة الفائقة-برشلونة، مزوّدَي التقانة الأوروبيين ARM، و Bull، و Gnodal، ومنظمات الحوسبة الفائقة الرئيسية المعنية، ومشروع PRACE (الشراكة للحوسبة المتقدمة في أوروبا Partnership for Advanced Computing in Europe) ويتضمن: Juelich، Leibniz- Rechenzentrum، أو LRZ، و GENCI، و CINECA. الغرض من المشروع نشر نظام HPC من الجيل الأول مبني من تقانات مضمّنة مُجدية في استعمال الطاقة، وإجراء البحوث اللازمة لتحقيق الأداء بمقياس إكسا مع تصاميم مُجدية في استعمال الطاقة.

يسعى مشروع DEEP بقيادة Forschungszentrum Juelich لتطوير منصة تمكين لمقياس إكسا واستمثال لمجموعة من رمازٍ كبير التحدي. تستند هذه المنظومة إلى عقودٍ سلمي وتصميم مسرّع-البنيان المقوي للعنقود-بمنزلة إثبات صحة مفهومٍ لمنظومة إنتاج PRACE ذي 100 بيتافلوب/ثانية. وإضافة إلى الشريك الرئيسي، Juelich، يشمل المشروع شركة إنتل، وشركة ParTec، ومركز LRZ، وجامعة هايدلبرغ، والمدرسة الألمانية لأبحاث علوم المحاكاة، وشركة Eurotech، ومركز الحوسبة الفائقة ببرشلونة، وشركة Mellanox، ومدرسة البوليتيكنيك الاتحادية في لوزان، وجامعة لوفين الكاثوليكية، والمركز الأوربي للبحوث والتكوين المتقدم للحسابات العلمية، ومعهد قبرص، وجامعة ريجنزبورغ، وتجمع CINECA، الذي يضم 70 جامعة في إيطاليا، والشركة العامة الجيوفيزيائية فيريتاس.

اليابان. في كانون الأول (ديسمبر) 2013، اختارت وزارة التربية والتعليم والثقافة والرياضة والعلوم والتقانة اليابانية (MEXT) شركة RIKEN لتطوير ونشر منظومة بمقياس إكسا بحلول عام 2020. ويعود اختيارها إلى تجربتها في تطوير وتشغيل الحاسوب K ذي 10 بيتافلوب/ثانية، الذي احتل مرتبة أسرع حاسوب فائق في العالم في عام 2011. المنظومة بمقياس إكسا هذه التي تُقدّر تكلفتها بـ 140 مليار "ين" ¥ (1.38 مليار دولار) سوف يكون تصميمها مستندًا إلى تركيبة من معالجاتٍ للأغراض العامة general-purpose processors ومسرّعات accelerators، وشركٍ ثلاثة بأبغى حواسيب رئيسيين يابانيين: فوجيتسو، هيتاشي، NEC، إضافة إلى الحصول على الدعم الفني من جامعة طوكيو، وجامعة تسوكوبا Tsukuba، ومعهد طوكيو للتقانة، وجامعة توهوكو Tohoku، وشركة RIKEN.

الصين. منظومة Tianhe-2 الصينية هي الحاسوب الفائق الأسرع في العالم اليوم. يحوي هذا الحاسوب 16,000 عقدة، تضم كل منها معالجي زيون من إنتل Intel Xeon وثلثاء معالجاتٍ مساعدة زيون فاي من إنتل Intel Xeon Phi coprocessors. ويضم أيضًا وصلاتٍ بينية interconnects عالية السرعة تملكية، تسمى TH Express-2، من تصميم الجامعة الوطنية لتقانة الدفاع (NUDT). تقود جامعة NUDT أبحاثًا حول المعالجات، والمترجمات (المصرفات)، والخوارزميات المتوازية، والمنظومات. واستنادًا إلى هذا العمل، يُتوقع أن تنتج الصين منظومات ذات 100 بيتافلوب/ثانية في عام 2016 مبنية كليًا من رقاقات صينية الصنع، وتحديدًا المعالج والوصلات البينية شن-واي Shen-Wei. وكان من المقرر ترقية منظومة Tianhe-2 من الذروة 55 بيتافلوب/ثانية إلى 100 بيتافلوب/ثانية في عام 2015، بيد أن وزارة التجارة الأمريكية قيّدت الصادرات من معالجات إنتل إلى كل من: جامعة NUDT، والمركز الوطني للحوسبة الفائقة في تشانغشا Changsha، والمركز الوطني للحوسبة الفائقة في غوانغ زهو Guangzhou، والمركز الوطني للحوسبة الفائقة في تيانجين Tianjin بسبب مخاوف الأمن القومي.

الولايات المتحدة. تاريخياً، امتد تشبيك الولايات المتحدة وبحوث تقانة المعلومات وبرنامج التنمية إلى عدة بعثات ووكالات بحثية، بقيادة رئيسية لوزارة الطاقة (DOE)، ووزارة الدفاع (DoD)، ومؤسسة العلوم الوطنية (NSF). إن وزارة الطاقة هي اليوم الأكثر نشاطاً على صعيد نشر منظومات الحوسبة العالية الأداء وتطوير خططٍ للحوسبة بمقياس إكسا. في المقابل، ركزت وزارة الدفاع ومؤسسة العلوم الوطنية أكثر على البنية الأساسية السيبرية وبحوث تقانات التمكين الواسعة، ومن ذلك الخدمات السحابية للبحوث وتحليلات المعطيات الكبيرة. وبرغم استمرار التخطيط، فإن الولايات المتحدة لم تعتمد بعد مبادرةً للحوسبة المتقدمة شبيهةً بتلك المبادرات التي تجري في أوروبا واليابان.

التعاون الدولي. ومع أن التنافس العالمي على قيادة الحوسبة المتقدمة وتحليلات المعطيات ما زال مستمرًا، فإن ثمة تعاونًا دوليًا فعالًا. المشروع الدولي للبرمجيات بمقياس إكسا (IESP: International Exascale Software Project) هو واحد من الأمثلة على ذلك في مجال الحوسبة المتقدمة. إن مشروع IESP، الذي حصل على تمويل أولي من الحكومات في اليابان، والاتحاد الأوربي، والولايات المتحدة، فضلاً على مساهماتٍ إضافية من الجهات المعنية في القطاع، وُضع للتمكين للميز الفائق الدقة والعلوم الكثيفة المعطيات والبحوث الهندسية حتى عام 2020.

بعد سلسلة من الاجتماعات، طُوّر فريق IESP الدولي خطة لبيئة حسابية عالية الجودة عالية ومشاركة لمنظومات بمقياسين: بيتا وإكسا (petascale/exascale). وسوف ينقل المنهج المرتبط بتطوير البرمجيات المجتمع من موقعه الحالي إلى الحوسبة بمقياس إكسا. [7]

الخلاصة

وصلت الحوسبة إلى نقطة انعطاف عميقة، اقتصادياً وتقنياً. إن نهاية تصدّد دينر Dennard scaling وأثاره في استمرار التقدم في تصميم أشباه النواقل، والتحول إلى الحوسبة النقالة والسحابية، والنمو الهائل للمعطيات العلمية، ومعطيات الأعمال، والمعطيات الحكومية، ومعطيات المستهلكين، والفرص المتاحة لتحليلات المعطيات والتعلم الآلي، والحاجة المستمرة إلى منظومات حوسبة قوية أكثر لدفع عجلة تقدّم العلوم والهندسة، كل ذلك يمثل إطار النقاش حول مستقبل الحوسبة بمقياس إكسا وتحليلات المعطيات الكبيرة. مع ذلك، ثمة أمور واضحة:

المعطيات الكبيرة ومقياس إكسا. إن تحليلات المعطيات الراقية (المعطيات الكبيرة) والحوسبة الراقية (مقياس إكسا) كلاهما عنصران أساسيان في جدول أعمال البحث والتطوير للحوسبة المتكاملة؛ ولا يجوز أن يُضحى أو يُقلل من شأن أي منهما لتطوير الآخر؛

الخوارزميات، والبرمجيات، والتطبيقات. إن البحث والتطوير في خوارزميات وبرمجيات، وتطبيقات الجيل القادم، لا يقل أهمية عن الاستثمار في تجهيزات أشباه النواقل والعتاديات؛ وتاريخياً لم يستثمر مجتمع البحوث إلى حدّ كافٍ في هذه المجالات؛

النظام البيئي لتقانة المعلومات (المعلوماتية). إن النظام البيئي العالمي لتقانة المعلومات في حالة تغير مستمر، مع التحول إلى جيل جديد من التجهيزات المحمولة ذات الطاقة المنخفضة، والخدمات السحابية، وتحليلات المعطيات الغنية؛

الأبحاث الخاصة والعالمية. إن منافسة القطاع الخاص والتعاون البحثي العالمي كلاهما ضروريان للتصدي لتصميم واختبار ونشر الحوسبة بمقياس إكسا ومقدرات تحليلات المعطيات.

وثمة فرص كبيرة وتحديات كبيرة في مجال الحوسبة المتقدمة، في الحساب وتحليلات المعطيات على حدٍ سواء. إنَّ الاكتشافات العلمية بواسطة العلوم الحاسوبية وتحليلات المعطيات هي حقاً "الحدود التي لا نهاية لها" التي تحدث عنها فانيفار بوش Vannevar Bush ببلاغة في عام 1945. إنَّ التحديات قائمة لجميع علوم الحاسوب للحفاظ على بحوث وتطوير ونشر البنية الأساسية للحوسبة العالية الأداء اللازمة للتمكين لتلك الاكتشافات.

عرفان

نحن شاكرون للرؤى ووجهات النظر التي تلقيناها من DARPA، ومن مجموعات دراسات العتاديات والبرمجيات والتطبيقات بمقياس إكسا في وزارة الطاقة DOE. ونعبّر عن عرفاننا أيضاً للتعليقات والاقتراحات الثاقبة من مراجعي المسودات السابقة من هذه المقالة. يُقرّر دانييل أريد Daniel A. Reed بالدعم الذي حصل عليه من مؤسسة العلوم الوطنية NSF تحت المنحة ACI-1349521. ويُقرّر جاك دونغارا بدعم المؤسسة نفسها تحت المنحة ACI-1339822 ودعم وزارة الطاقة تحت منحة وزارة الطاقة DEFG02-13ER26151.

المراجع

- [1] Amarasinghe, S. et al. *Exascale Software Study: Software Challenges in Extreme-Scale Systems*. Defense Advanced Research Projects Agency, Arlington, VA, 2009; <http://www.cs.rice.edu/~vs3/PDF/Sarkar-ACS-July-2011-v2.pdf>
- [2] American Association for the Advancement of Science. *Guide to R&D Funding - Historical Data*. AAAS, Washington, D.C., 2015; <http://www.aaas.org/page/historical-trends-federal-rd>
- [3] Chang, F. et al. Bigtable: A distributed storage system for structured data. *ACM Transactions on Computer Systems* 26, 2 (June 2008), 4:1–4:26.
- [4] Datta, K. et al. Stencil computation optimization and auto-tuning on state-of-the-art multicore architectures. In *Proceedings of the 2008 ACM/IEEE Conference on Supercomputing* (Austin, TX, Nov. 15–21). IEEE Press, Piscataway, NJ, 2008, 1–12.
- [5] Dennard, R.H., Gaensslen, F.H., Yu, H.-n., Rideout, V.L., Bassous, E., and LeBlanc, A.R. Design of ion-implanted MOSFETs with very small physical dimensions. *IEEE Journal of Solid State Circuits* 9, 5 (Jan. 1974), 256–268.
- [6] Dongarra, J.J. The LINPACK benchmark: An explanation. In *Proceedings of the First International Conference on Supercomputing* (Athens, Greece, June 8–12). Springer-Verlag, New York, 1988, 456–474.
- [7] Dongarra, J.J. et al. The international exascale software project roadmap. *International Journal of High Performance Computing Applications* 25, 1 (Feb. 2011), 3–60.
- [8] Esmailzadeh, H., Blem, E., Amant, R.S., Sankaralingam, K., and Burger, D. Dark silicon and the end of multicore scaling. In *Proceedings of the 38th Annual International Symposium on Computer Architecture* (San Jose, CA, June 4–8). ACM, New York, 2011, 365–376.
- [9] Fuller, S.H. and Millett, L.I. Computing performance: Game over or next level? *Computer* 44, 1 (Jan. 2011), 31–38.
- [10] Geist, A. and Lucas, R. Major computer science challenges at exascale. *International Journal of High Performance Applications* 23, 4 (Nov. 2009), 427–436.
- [11] Gill, P., Jain, N., and Nagappan, N. Understanding network failures in data centers: Measurement, analysis, and implications. *Proceedings of ACM SIGCOMM* 41, 4 (Aug, 2011), 350–361.

- [12] Hey, T., Tansley, S., and Tolle, K. *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft Research, Redmond, WA, 2009; http://research.microsoft.com/en-us/UM/redmond/about/collaboration/fourthparadigm/4th_PARADIGM_BOOK_complete_HR.pdf
- [13] Kamil, S., Shalf, J., and Strohmaier, E. Power efficiency in high-performance computing. In *Proceedings of the Fourth Workshop on High-Performance, Power-Aware Computing* (Miami, FL, Apr.). IEEE Press, 2008.
- [14] Kogge, P., Bergman, K., Borkar, S. et al. *Exascale Computing Study: Technology Challenges in Achieving Exascale Systems*. U.S. Defense Advanced Research Projects Agency, Arlington, VA, 2008; <http://www.cse.nd.edu/Reports/2008/TR-2008-13.pdf>
- [15] Lucas, R., Ang, J., Bergman, K., Borkar, S. et al. *Top Ten Exascale Research Challenges*. Office of Science, U.S. Department of Energy, Washington, D.C., Feb. 2014; <http://science.energy.gov/~media/ascr/ascac/pdf/meetings/20140210/Top10reportFEB14.pdf>
- [16] Meuer, H., Strohmaier, E., Dongarra, J. and Simon, H. *Top 500 Supercomputer Sites, 2015*; <http://www.top500.org>
- [17] Nobelprize.org. Nobel Prize in Chemistry 2013; http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2013/press.html
- [18] Olston, C., Reed, B., Strivastava, U., Kumar, R., and Tomkins, A. Pig Latin: A not-so-foreign language for data processing. In *Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data* (Vancouver, BC, Canada, June 9–12). ACM Press, New York, 2008, 1099–1110.
- [19] Partnership for Advanced Computing in Europe (PRACE). 2014; <http://www.prace-ri.eu/>
- [20] Pinheiro, E., Weber, W.-D., and Barroso, L.A. Failure trends in a large disk drive population. In *Proceedings of the Fifth USENIX Conference on File and Storage Technologies* (San Jose, CA, Feb. 13–16). USENIX Association, Berkeley, CA, 2007.
- [21] Schroeder, B., Pinheiro, E., and Weber, W.-D. DRAM errors in the wild: A large-scale field study. In *Proceedings of the 11th International Joint Conference on Measurement and Modeling of Computer Systems* (Seattle, WA, June).
- [22] Schroeder, B., Pinheiro, E., and Weber, W.-D. DRAM errors in the wild: A large-scale field study. In *Proceedings of the 11th International Joint Conference on Measurement and Modeling of Computer Systems* (Seattle, WA, June). ACM Press, New York, 2009, 193–204.
- [23] U.S. Department of Energy. *Synergistic Challenges in Data-Intensive Science and Exascale Computing*. Report of the Advanced Scientific Computing Advisory Committee Subcommittee, Mar. 30, 2013; http://science.energy.gov/~media/ascr/ascac/pdf/reports/2013/ASCAC_Data_Intensive_Computing_report_final.pdf
- [24] U.S. Department of Energy. *The Opportunities and Challenges of Exascale Computing*. Office of Science, Washington, D.C., 2010; http://science.energy.gov/~media/ascr/ascac/pdf/reports/Exascale_subcommittee_report.pdf
- [25] White, T. *Hadoop: The Definitive Guide*. O'Reilly Media, May 2012.

الكاتبان

دانييل أ. ريد (dan-reed@uiowa.edu) هو نائب الرئيس للبحوث والتنمية الاقتصادية، وأستاذ علوم الحاسوب وهندسة الكهرباء والحوسيب والطب، في جامعة ايوا Iowa.

جاك دونغارا (dongarra@icl.utk.edu) حاصل على منصب في جامعة تينيسي Tennessee، ومختبر أوك ريدج الوطني، وجامعة مانثستر.

المعطيات الكبيرة وتحدياتها التقنية

BIG DATA AND ITS TECHNICAL CHALLENGES*

H.V. Jagadish, J. Gehrke, A. Labrinidis, Y. Papakonstantinou, J. M. Patel, R. Ramakrishnan, C. Shahabi

ترجمة: د. محمد عباسي
مراجعة: د. أحمد حصري

عرض للتحديات التقنية الجوهرية لإدراك إمكانات المعطيات الكبيرة.

تُجمَع المعطيات، في طيف واسع من التطبيقات، على نحو غير مسبوق. وقد أصبحت القرارات، التي كانت سابقاً تعتمد على التخمين، أو على نماذج الواقع المنفذة يدوياً، تُتخذ الآن اعتماداً على نماذج رياضية قوامها المعطيات. يتحكم تحليل المعطيات الكبيرة هذا، بكافة مناحي المجتمع، ويدخل فيها خدمات الهواتف النقالة، وتجارة التجزئة، والتصنيع، والخدمات المالية، وعلوم الحياة، والعلوم الفيزيائية.

لندرس، على سبيل المثال بحثاً علمياً جرى تطويره بمعطيات كبيرة^{1,12}. لقد نَقَلَ استقصاء السماء الرقمي "سلون"²³ 1 علم الفلك من حقلٍ كان النقاط صور السماء فيه يُمتلَّ جزءاً كبيراً من عمل عالم الفلك، إلى حقلٍ تُجمَع فيه الصور الفضائية في قاعدة معطيات، ويقتصر عمل عالم الفلك على إيجاد أشياء وظواهر مميزة باستعمال قاعدة المعطيات. في علوم الأحياء، ثمة تقليد راسخ يتمثل في إيداع المعطيات العلمية في مخزن عام، وفي إنشاء قواعد معطيات عامة يستعملها علماء آخرون. إضافة إلى ذلك، ومع تطور التكنولوجيا، خاصة مع اختراع "تسلسل الجيل القادم (NGS)"، بدأ عدد مجموعات المعطيات التجريبية المتاحة يزداد أسياً.

مفاهيم مفتاحية

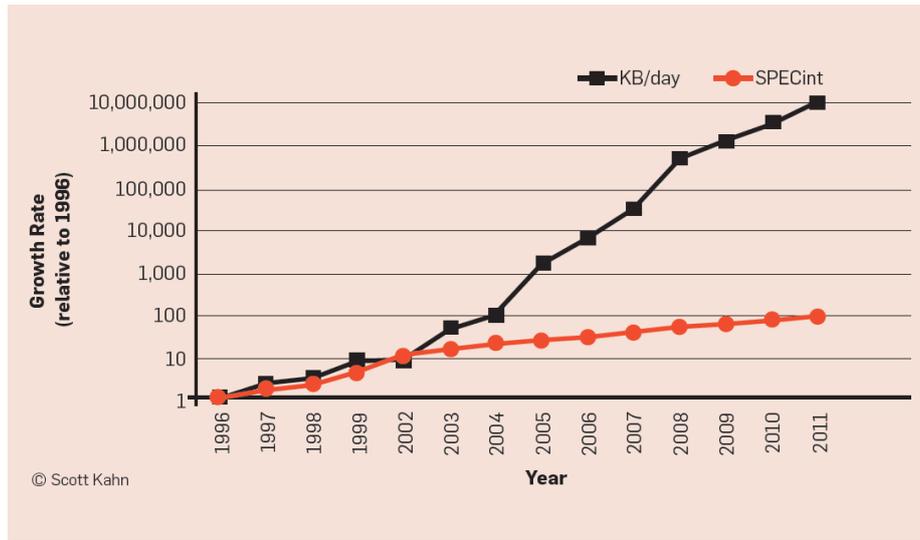
- تُحدث المعطيات الكبيرة ثورةً في جميع مناحي حياتنا بدءاً من الشركات وانتهاءً بالمستهلكين، ومن العلم إلى الحكومة.
- إن بناء قيمةٍ من المعطيات الكبيرة عمليةٌ متعددة المراحل تشمل على: التحصيل، واستخلاص المعلومات وتصفيته، ومكاملة المعطيات، والنمذجة والتحليل، والتفسير والاستعمال. يُركِّز كثير من النقاشات التي تدور حول المعطيات الكبيرة على خطوةٍ أو اثنتين، متجاهلة البقية.
- ازدادت تحديات البحث، بدءاً من تباين المعطيات، إلى عدم الترابط وعدم الاكتمال، وإلى الديمومة، وإلى الخصوصية، وإلى التصوُّر والمشاركة، وإلى بيئة الأدوات في محيط المعطيات الكبيرة.
- يُظهر كثير من دراسات الحالة، أن ثمة مكافآت ضخمة بانتظار أولئك الذين يستعملون المعطيات الكبيرة على نحو صحيح

* تُشير هذا البحث في مجلة *Communications of the ACM*، المجلد 57، العدد 7، تموز (يوليه) 2014، الصفحات 86 – 94.

¹ Sloan Digital Sky Survey نظام تصوير واستقصاء للفضاء يستعمل تلسكوباً مزوداً بعدسة قطرها 2.4 م، وهو موضوع في مرصد بولاية نيو مكسيكو في الولايات المتحدة الأمريكية. (المترجم)

يبين الشكل 1 معدل نمو (growth rate) الناتج لطُرق NGS² الحالية مُعبراً عنه بتسلسل المعطيات الخام الناتج عن آلة NGS مفردة، كما يُبين زيادة أداء معيار وحدة المعالجة المركزية SPECint. من الواضح أن نمو معطيات التسلسل NGS يتخطى أرباح الأداء التي قدمها قانون مور لتطبيقات النظم المفردة هنا (SPECint). مع ملاحظة أن حجم معطيات التسلسل في الشكل 1 هو ناتج تحليل الصور الخام التي أنتجتها فعلياً آلات NGS. إن حجم مجموعات معطيات الصور الخام هذه كبير جداً (يُناهز التيرا³ بايت لكل مخبر في اليوم) بحيث يغدو التفكير في تخزينها غير عملي. في الواقع يجري تحليل هذه الصور إبان الرحلة لإنتاج تسلسل معطيات، يُحفظ فيما بعد.

للمعطيات الكبيرة القدرة على إحداث ثورة تتخطى كثيراً مجال البحث. فقد أدى عمل محرّك البحث غوغل على نظام الملفات ونظام تخفيض الخرائط (MapReduce)⁴، ومن ثمّ باعتماد المصدر المفتوح على نُظم من قبيل نظام Hadoop⁵، إلى تطوير واعتماد تقنيات المعطيات الكبيرة على نحو شامل، من قبيل شركات تُركّز على الوب، من قبيل فيسبوك، ولينكدان LinkedIn، ومايكروسوفت، وكوانت كاست Quantcast، وتويتر، وياهو! غدت هذه التقنيات الأساس الذي لا غنى عنه في تطبيقات تمتد من البحث في الوب إلى اقتراح المحتوى والإعلان الحاسوبي. ثمة حالات جرى عرضها تُبين



الشكل 1. حجم معطيات تسلسل الجيل القادم NGS مقارنة بـ SPECint.

² Next Generation Sequencing: تسلسل الجيل القادم أو (التسلسل المتوازي الكثيف)، في علم الأحياء، يدل على مجموعة من تقنيات سلسلة الدنا (DNA) يمكنها سلسلة الدنا على نحو ضخم (بمقياس الجيغا). حلّت هذه الطرق مكان طُرق السلسلة التقليدية التي استمرّت أكثر من أربعين عاماً. وجرى تطوير وتسويق عدة منصات مختلفة تعتمد هذه التقنيات. (المترجم)

³ Tera: تيرا - إحدى البادئات الممثلة لقوى العشرة، والتي توضع أمام اسم الوحدة المستعملة. رمزها T وتمثل 10^{12} . فمثلاً Tj تعني 10^{12} جول أي تيرا جول. (المترجم)

⁴ MapReduce: نموذج برمجة لمعالجة مجموعات كبيرة من المعطيات وتوليدها بخوارزمية متوازية وموزعة. (المترجم)

⁵ Hadoop: إطار عمل برمجة يعتمد اللغة جافا، ويدعم معالجة مجموعات كبيرة من المعطيات في بيئة حوسبة موزعة. (المترجم)

قيمة المعطيات الكبيرة في التأمين الصحي (وذلك بالمراقبة المنزلية المستمرة، وبالمكاملة بواسطة المزودين)،³ وفي التخطيط الحضري (بدمج المعطيات الجيولوجية العالية الدقة)، وفي حركة النقل الذكي (بتحليل وتمثيل معطيات شبكة الطرقات التفصيلية)، وفي نمذجة البيئة (باستعمال المعطيات المجمعّة لشبكات الحساسات المنتشرة في كل مكان)،⁴ وفي حفظ الطاقة (ببيان أنماط الاستعمال¹⁸)، وفي المواد الذكية (بالاستفادة من مبادرة جينات المواد الحديثة)، وفي الترجمة الآلية بين اللغات الطبيعية (بتحليل المجاميع الكبيرة)، وفي التعليم (وعلى وجه الخصوص في الدورات التعليمية المباشرة على الإنترنت)،² وفي علوم المجتمع (باعتقاد منهجية جديدة تزداد انتشاراً بسرعة بسبب تكلفة تحصيل المعطيات المنخفضة جداً)،¹⁴ وفي تحليل المجازفة التّسقي في الموارد المالية (بمكاملة تحاليل شبكة العقود لإيجاد التبعية بين الكيانات المالية)،⁸ وفي الأمن الوطني (بتحليل الشبكات الاجتماعية والتحويلات المالية للإرهابيين المحتملين)، وفي أمن الحاسوب (بتحليل وقائع تسجيل الدخول، المعروفة باسم معلومات الأمن وإدارة الأحداث، أو اختصاراً SIEM)، وغيرها.

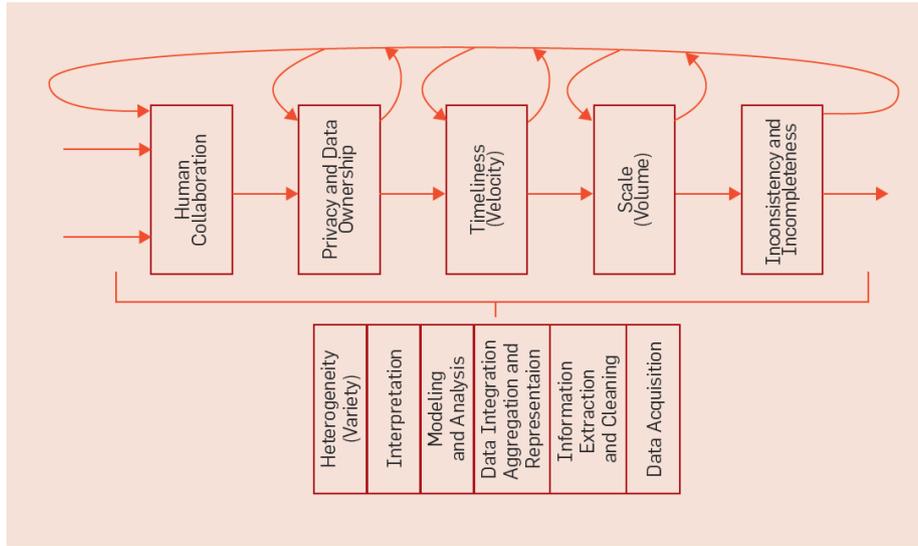
في عام 2010 خزنت شركاتٌ ومستخدمون أكثر من 13 إكسا⁶ بايت من معطيات جديدة؛ وهذا يفوق بأكثر من خمسين ألف ضعف حجم المعطيات في مكتبة الكونغرس. فُدرت القيمة المُحتَملة لمعطيات نظام تحديد الموقع الشخصي العالمي بـ 700 بليون دولار للمستخدمين النهائيين، ويمكن أن تؤدي، استناداً إلى تقرير حديث لماكينزي،¹⁷ إلى تخفيض ما يقارب 50% من تكلفة تطوير المُنتج وتجميعه. وتتبعاً لماكينزي⁷ بتأثير ضخم مماثل للمعطيات الكبيرة في مجال التوظيف، حيث ستحتاج الولايات المتحدة ما بين 140000 و 190000 عاملٍ ذي خبرة "تحليلية عميقة"؛ كما ستحتاج إلى تأهيل 1.5 مليون مدير في مجال المعطيات. وليس مدهشاً أن يُصدر مجلس مستشاري الرئيس الأمريكي للعلوم والتكنولوجيا حديثاً تقريراً عن التشبيك وعن البحث والتطوير في مجال تقنية المعلومات²² يُعرّف المعطيات الكبيرة على أنها " حقل بحث جديد " يُمكن أن يُسرّع " التطور في طيف واسع من الأولويات ". كما أن وسائل الإعلام الرائجة تُقدّر الآن قيمة المعطيات الكبيرة، وهذا واضح من معالجة هذا الموضوع في صحف مثل؛ الإكونومست،⁷ والنيويورك تايمز،^{15,16} والناشيونال بابلوك راديو،^{19,20} ومجلة فوربس⁹.

ومع أن الفوائد الكامنة للمعطيات الكبيرة حقيقية وهامة، وحقق بعضها نجاحاً أولياً (مثل نظام "سلون" لمسح السماء الرقمي)، فإن ثمة تحديات تقنية كثيرة مازالت متبقية لابدأ من مواجهتها لتحقيق الفوائد تحقيقاً كاملاً. يعتبر حجم المعطيات المجرّد تحدياً رئيسياً، وهو الأسهل تعريفاً. ومع ذلك فثمة تحديات أخرى قائمة. تميل شركات تحليل الصناعة إلى بيان أن التحديات ليست في الحجم فحسب، بل في التنوع والسرعة¹⁰ أيضاً، وأنّ على الشركات أن تركز عليها جميعاً. يدلّ التنوع على تباين أنواع المعطيات، وعلى التمثيل والتأويل الدلالي. على حين تدلّ السرعة على معدّل وصول المعطيات وعلى الإطار الزمني الذي ستعمل ضمنه. ومع أن هذه العناصر الثلاثة هامة، فإن هذه القائمة المختزلة لا تتضمن متطلبات هامة إضافية. وقد اقترحت جهات متعددة عدة متطلبات إضافية، مثل صحة المعطيات وسلامتها. وتبقى أيضاً قضايا مهمة من مثل الخصوصية وقابلية الاستعمال.

إن تحليل المعطيات الكبيرة عملية تكرارية، ولكل منها تحدياته الخاصة، ويتضمن كثيراً من المراحل المُتميزة كما يظهر في الشكل 2. ندرس هنا دورة حياة المعطيات الكبيرة من البداية حتى النهاية.

⁶ Exa: بادئة تمثّل 10¹⁸. (المترجم)

⁷ McKinsey: شركة أمريكية استشارية متعددة الجنسيات تهتم بقضايا الإدارة، وتُجري تحليلات كمية ونوعية لتقييم قرارات الإدارة. (المترجم)



الشكل 2. مخطط سير تحليل المعطيات الكبيرة. يُظهر النصف العلوي من الشكل الخطوات الرئيسية في تحليل المعطيات الكبيرة. لاحظ حلقات التغذية العكسية الممكنة في كافة المراحل. يبين النصف السفلي من الشكل خصائص المعطيات الكبيرة التي تجعل هذه المراحل تحدياً.

مراحل دورة حياة المعطيات الكبيرة

يُرَكِّز كثير من الناس على خطوة التحليل/النمذجة وحسب- ومع أن هذه الخطوة جوهرية، إلا أن فائدتها قليلة بدون المراحل الأخرى لسير تحليل المعطيات. فمثلاً، علينا مقارنة السؤال: أي المعطيات يجب أن نُسجّلها، باعتبارها قيمة، ربما بطرق لا يمكننا توقعها تماماً، وأن نُطوّر طُرُقاً لاستنتاج قيم من معطيات أُخذت على نحو غير كامل أو غير صحيح. يوُلِّد ذلك الحاجة إلى تَعَقُّب المصدر وإلى معالجة الخطأ والالتباس. وكمثال آخر، عندما تُمَثَّل ذات المعلومات على نحو مكرر ومتشابه، فإن ذلك يسمح لنا باستعمال التقنيات الإحصائية لمواجهة تحديات من قبيل مكاملة المعطيات واستخلاص الكيان/العلاقة. وهذا هو مفتاح الاستخلاص الناجح لمعطيات أُخذت من مصادر متعددة (مثلاً، تجارب متشابهة نُشرت نتائجها من مخابر مختلفة، معلومات عن حركة السير متعددة المصادر، معطيات عن حفل مُعَيَّن من قبيل وسائل الترفيه، مُختارة من مواقع وب مختلفة). هذه العناوين أساسية للنجاح، ومع هذا فقلماً تُذكر بنفس الرّحم الذي تذكر فيه المعطيات الكبيرة. وحتى في مرحلة التحليل، التي جلبت كثيرا من الانتباه، ثمة تعقيدات غير مفهومة في سياق التجمعات المتعددة الملكية حيث تُنفَّذ برامج عدة مستخدمين في الوقت ذاته.

سنبدأ، فيما تبقى من هذه المقالة، بدراسة المراحل الخمس في سير المعطيات الكبيرة، مع التحديات الخاصة بكل مرحلة. وسنقدّم أيضاً دراسة حالة بوصفها مثالاً للقضايا التي تبرز في المراحل المختلفة. والتحديات المُتشابهة السّنة المتعلقة بها.

تحصيل المعطيات. لا تأتي المعطيات الكبيرة من الفراغ: إنها تسجيل لنشاط أساسي هام. لندرس، على سبيل المثال، قدرتنا على استشعار العالم من حولنا، ومراقبته بدءاً من معدل ضربات قلب مواطن عجوز، إلى وجود السموم في الهواء الذي نستنشق، إلى تسجيلات نشاطات المستخدمين على موقع وب أو قيّد الوقائع في نظام برمجي. تُنتج المُحسّات،

دراسة حالة

منذ خريف العام 2010 ، وكجزء من عقد مع سلطة النقل المحليّة في لوس أنجلوس (LA-Metro) سُمح لباحثين من جامعة جنوب كاليفورنيا عاملين في مركز نُظُم الإعلام المتكاملة (IMSC) بالنفاذ إلى معطيات عن النقل مكانية وزمنية عالية الدقة من شبكة طُرُق مقاطعة لوس أنجلوس. تبلغ سرعة ورود هذه المعطيات 46 ميغا بايت في الدقيقة، وُجِع إلى الآن ما يزيد عن 15 تيرا بايت. طور الباحثون في مركز نُظُم الإعلام المتكاملة نظاماً كاملاً أسموه اختصاراً ترانس دِك TransDec، لآخذ قرارات النقل، لتحصيل وتخزين وتحليل وعرض مجموعات المعطيات هذه (انظر الشكل المرفق). سنناقش هنا الأجزاء المختلفة للمشروع ترانس دِك التي تقابل مخطط سير المعطيات الكبيرة الموضَّح في الشكل 2.

التحصيل: يحصل النظام الحالي مجموعات المعطيات التالية في الزمن الحقيقي:

- **كواشف حلقة المرور:** يجمع ما يقارب 8900 مُجس، موضوع على الطرقات السريعة والشوارع الرئيسية، معاملات المرور مثل الانشغالية، والحجم والسرعة بمعدل قراءة واحدة لكل مُجس كل دقيقة.
- **الحافلات والقطارات:** متضمنة معلومات عما يقارب 2036 حافلة و 35 قطاراً تعمل على 145 خطاً مختلفاً في مقاطعة لوس أنجلوس. تحتوي معطيات المُجس موضع كل حافلة زمنياً ومكانياً، مرة كل دقيقتين، ومعلومات الموقف التالي نسبة إلى الموقع الحالي، ومعلومات عن التأخر مقارنة بالجدول الزمنية الموضوع سلفاً.
- **عدادات التحكم في المرور والإشارات المتغيرة المحتوى:** ثمة 1851 عداد تحكّم في المرور¹ تُنظّم حركة المرور للحافلات الداخلة إلى الطرقات السريعة حسب حالات المرور، و 160 إشارة متغيرة المحتوى² لتزويد المسافرين بمعلومات عن حالات الطريق من قبيل التأخيرات، والحوادث، ومناطق الأشغال على الطرقات. تُحدّث معلومات كل عداد تحكّم وكل إشارة متغيرة المحتوى كل 75 ثانية.
- **الأحداث:** معلومات نصية مُفصّلة (مثل، عدد المصابين، وزمن وصول سيارة الإسعاف) عن الأحداث الخاصة التي من قبيل التصادمات، مخاطر المرور، وغيرها، مُحصّلة من ثلاث وكالات مختلفة.

التنظيف: تزيل خوارزميات تنقية المعطيات عناوين XML الفائضة، كما تكشف قراءات المُحسّات الفائضة وتزيلها، وما شابه، في الزمن الحقيقي باستعمال تطبيق Stream Insight من مايكروسوفت، وهذا ما يؤدي إلى تخفيض معدل الدخل من 65 ميغا بايت في الدقيقة إلى 25 ميغا بايت في الدقيقة. تُصخ النتائج بعد ذلك على شكل جداول بسيطة إلى المنصة السحابية أزور Microsoft Azure.

التجميع والتمثيل: تُجمع المعطيات وتُصنّف في مجموعة من الجداول بلغة أوراكل 11g (مبوّية بالمكان وبالزمن، وباستعمال الشجرة-R والشجرة-B). على سبيل المثال، تُجمع المعطيات لإنشاء مخططات لدعم مجموعة محددة سلفاً من الاستعلامات المكانية والزمنية (مثلاً، معدل السرعة الساعية لأحد المقاطع في منطقة نورث بوند I-110).

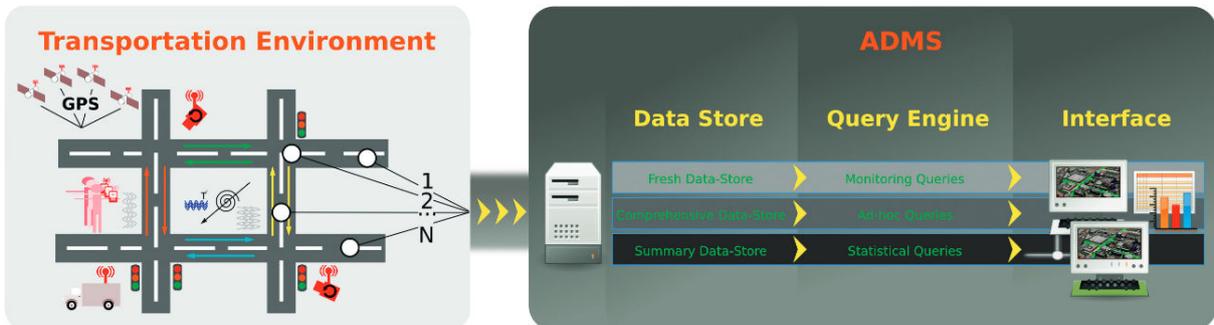
التحليل: تُطبّق عدة تقنيات تُعلّم آلة لتوليد نماذج/أنماط دقيقة لمقاطع طُرُق متنوعة في مقاطعة لوس أنجلوس في أوقات مختلفة من اليوم (مثلاً، ساعات الذروة)، وفي أيام مختلفة من الأسبوع (مثلاً، نهاية الأسبوع)، وفي الفصول المختلفة. تُستعمل معطيات الحوادث القديمة لتصنيف الحوادث الجديدة وذلك للتنبؤ بزمن الإخلاء وبمدة تعطل المرور الحاصل.

التفسير: قد تحدث أخطاء كثيرة في نظام مُعدّد، مما يؤدي إلى نتائج زائفة. فعلى سبيل المثال، قد تتعطل عناصر (مُستقلة) مختلفة من النظام دون ملاحظة ذلك، مما يؤدي إلى فقد في المعطيات. بالمقابل قد تُغيّر إحدى الشركات مصاعغة المعطيات دون إعلام شركة نقل المعطيات، وهذا ما يؤدي إلى تحليل خاطئ. ولمواجهة مثل تلك المسائل، جرى تطوير عدة خطط مراقبة، على التوازي مع آليات للحصول على تأكيدات المُستخدم وتصحيحاته.

¹ Ramp meter : جهاز إشارة مرور ثنائي الأضواء (بدون لون برتقالي) مع نظام تحكّم. (المترجم)

² Changeable Message Signs (CMS): لوحات إظهار مرورية يمكن التحكم في الرسائل أو التعليمات التي تظهر عليها. (المترجم)

TransDec.



© Luciano Nocera

وعمليات المحاكاة، والتجارب العلمية، أحجاماً كبيرة من المعطيات هذه الأيام. فعلى سبيل المثال سينتج المقرب المصفوفي، المُخطَّط بناؤه على مساحة كيلو متر مربع، مليون تيرا بايت من المعطيات الخام في اليوم.

يمكن تصفية وضغط الكثير من هذه المعطيات عدة مراتب كَبْر دون الإخلال بالقدرة على استخلاص النشاط الأساسي الهام. يتمثل أحد التحديات، في تعريف هذه المرشحات " المتصلة مباشرة " بطريقة لا تُهمل فيها معلومات هامة، إذ إن المعطيات الخام ضخمة جداً بحيث لا يمكن فرزها جميعاً. فالمعطيات المُجمَّعة من مُحسَّات، على سبيل المثال، عادة ما تكون مترابطة مكانياً وزمنياً (مثل مُحسَّات حركة السير على ذات المقطع من الطريق). لنفرض أن قراءة أحد المُحسَّات تختلف عن قراءة باقي المُحسَّات على نحو كبير جداً. من المُرجَّح أن يكون ذلك ناتجاً عن خلل في المُحسِّن، ولكن كيف يمكننا التحقق أن هذه القراءة ليست ذات دلالة حقيقية؟

إضافة إلى ذلك، يُعتبر إساق مجموعات معطيات كبيرة تحدياً هاماً، خاصة إذا ما رافق ذلك عمليات تصفية معطيات ونقلها مباشرة، ومن ثم علينا استعمال تقنيات ازرداد⁸ تزايدية فعالة. وقد لا يكون ذلك كافياً لكثير من التطبيقات، ومن ثم يجب تصميم معالجة موضعية فعالة.

استخلاص المعلومات وتنقيتها. لا تكون المعلومات المجموعة ذات مصاغة جاهزة للتحليل غالباً. فعلى سبيل المثال، لتأمل مجموعة السجلات الصحية الإلكترونية في مستشفى، المؤلف من إملاءات مدونة من عدة أطباء، ومعطيات منظمّة من مُحسَّات، ومن قياسات (ربما مع بعض الغموض المُرافق)، ومعطيات صورة، بأشعة سينية مثلاً، ومقاطع فيديو من مسابر. لا يمكننا هنا إبقاء المعطيات بهذه الصيغة وتحليلها على نحو فعال. بل نحتاج إلى عملية استخلاص تأخذ المعلومات المطلوبة من المصادر الأساسية وتُعبّر عنها بصيغة منظمّة مناسبة للتحليل. يُعتبر القيام بهذه العملية على وجه صحيح وكامل تحدياً تقنياً مستمراً. تعتمد عملية الاستخلاص هذه غالباً على التطبيق (فما نحتاج استخراجه من المرنان المغناطيسي MRI مختلف كلياً عما نحتاج استخراجه من صورة للنجوم، أو من لقطة مراقبة). تتطلّب قضايا الإنتاجية اعتماد طرق توضيحية توصف بدقة أهداف استخلاص المعلومات، ومن ثم استمثال تنفيذ هذه المهام عند معالجة معطيات جديدة.

إن معظم مصادر المعطيات غير موثوقة: فالمُحسَّات يمكن أن تكون غير مضبوطة، والبشر يمكن أن يُقدّموا آراءً مُنحازة، ومواقع الوب البعيدة يمكن أن تكون مفقودة أو خارج الصلاحية، وهكذا. يُعتبر إدراك مصادر الخطأ هذه ونمذجتها، خطوة أولى نحو تطوير تقنيات تنقية المعطيات. ولسوء الحظ، فإن معظم هذا العمل يعتمد كثيراً على مصدر المعلومات وعلى التطبيق.

مكاملة المعطيات، وجمعها، وتمثيلها. يتطلّب التحليل الواسع الفعال غالباً تجميع معطيات غير متجانسة من مصادر متعددة. فمثلاً للحصول على نظرة صحية شاملة لمريض (أو لمجتمع) لا بد من مكاملة سجل المريض الطبي وتحليله مع المعطيات البيئية المتاحة من الإنترنت، ومع قراءات من أنواع مختلفة من المقاييس (على سبيل المثال، مقاييس السكر، ومقاييس ضربات القلب، ومقاييس التسارع، وغيرها³). تساعد مجموعة أدوات تحويل المعطيات ومكاملتها محلّ المعطيات على حلّ مسألة عدم التجانس في بنية المعطيات ودلالاتها. يُفرضي حلّ مسألة عدم التجانس إلى معطيات مكاملة قابلة للتفسير ضمن مجتمع معين، إذ إنها تطابق مخططاته المعيارية ومتطلبات التحليل. إلا أن تكلفة المكاملة

⁸ الازرداد (data ingestion): الحصول على المعطيات ونقلها ومعالجتها تمهيداً للاستعمال أو التخزين في قاعدة معطيات. وتتضمن العملية تغيير الملفات على نحو منفرد بتعديل محتوياتها و/أو مصاغاتنا لاستيعابها ضمن ملف أكبر. (الترجم)

الشاملة غالباً مُرعبة، كما أن متطلبات التحليل تتغير سريعاً، ولهذا توفّر تقنيات المُكاملة الحديثة العهد "ادفع كلما مضيت pay-as-you-go" انفراجاً مغريباً، وذلك بإجراء معظم العمل بسرعة دعماً للاستقصاء المخصص لهذا الغرض. من الملاحظ أن الوفرة الهائلة للمعطيات على الإنترنت، مُقترنة بأدوات التحليل والمُكاملة، تؤدي إلى إنتاج معطيات مُستخلصة، تقود إلى نوع آخر من تزايد المعطيات، لا يتعلّق فقط بمشكلة حجم المعطيات، بل يتعداها إلى مشكلة تَعَقُّب مصدر هذه المعطيات المُستخلصة (كما سنوضّح لاحقاً).

ثمة الكثير من الطرق البديلة لتخزين المعلومات نفسها حتى لأبسط أنواع التحليل المعتمد على مجموعة معطيات واحدة، ولكل من هذه الطرق خياراته التفضيلية الخاصة. لاحظ، على سبيل المثال، التنوع الهائل في بنية مجموعات معطيات المعلومات البيولوجية العائدة لكيانات شديدة التشابه، من قبيل الجينات. يُعتبر تصميم قاعدة المعطيات اليوم فناً يُنفّد بعناية وفق مفهوم الشركة من قِبَل مختصين ذوي أجور مرتفعة. وعلينا أن نُمكن مختصين آخرين، من قبيل علماء ضمن هذا المجال، من إنشاء مخازن معطيات فعّالة، إما بواسطة أدوات تقسيم تساعد في عملية التصميم، أو بتجاوز عملية التصميم كلياً وتطوير تقنيات تُمكن من استعمال مجموعات المعطيات بفاعلية في غياب تصميم قاعدة معطيات ذكية.

النمذجة والتحليل. تختلف طُرُق استعمال واستخراج المعطيات الكبيرة جذرياً عن طرق التحليل الإحصائي التقليدية التي تُجرى على عينات صغيرة. فالمعطيات الكبيرة عادة يشوبها شغب، وهي متقلّبة، وغير متجانسة، وذات علاقة بيئية، وغير موثوقة. ومع ذلك، يمكن للمعطيات الكبيرة التي يشوبها شغب أن تكون أكبر قيمة من العينات الصغيرة، لأن الإحصاءات العامة التي نحصل عليها من أنماط منكرة ومن تحليل الترابط تفوق عادة التغيرات الفردية، وغالباً ما تكشف معرفةً وأنماطاً مخفية موثوقة أكثر. في الحقيقة، ومع اهتمام مناسب بالإحصاء، يمكن استعمال التحليل التقريبي للحصول على نتائج جيدة دون أن يُرهق التحليل بالحجم.

التفسير. في نهاية المطاف، على مُتخذ القرار، المُزوّد بنتيجة التحليل، أن يُفسّر هذه النتائج. وهذا يتضمن فحص كل الافتراضات الموضوعية وتتبع التحليل مرة ثانية. إضافة إلى ذلك، ثمة عدة مصادر مُحتملة للخطأ: فقد تحتوي النظم الحاسوبية على عثرات، وللنماذج دائماً تقريباً افتراضات، ويمكن أن تُبنى النتائج على معطيات خاطئة. لكل هذه الأسباب، لا يُسلّم أي مُستخدم مسؤول دفة القيادة إلى نظام حاسوب. بل يحاول أن يفهم النتائج التي أنتجها الحاسوب، وأن يتحقق منها. وعلى الحاسوب أن يُسهّل المهمة له. ويُعدّ هذا تحدياً عند التعامل مع المعطيات الكبيرة بسبب تعقيداتها. إذ إن ثمة افتراضات جوهرية خلف المعطيات المُسجّلة. كما تشمل الطُرُق التحليلية خطوات متعددة، مع افتراضات مُضمّنة أيضاً. وقد أكّدت الصدمة التي أصابت النظام المالي، والمتعلقة بقروض السكن، الحاجة لمثل هذا الاجتهاد في اتخاذ القرار - فعوضاً عن قبول الملاءة المُقرّرة من هيئة مالية، فإن على مُتخذ القرار أن يفحص بعناية الافتراضات الكثيرة في مراحل متعددة من التحليل. باختصار، ليس كافياً فقط أن نُزوّد بالنتائج، بل يجب تزويد المستخدمين بالقدرة على تفسير النتائج التحليلية المُستخرجة، وعلى إعادة التحليل بافتراضات، أو بمُعاملات، أو بمجموعات معطيات مختلفة لدعم طريقة التفكير البشري والظروف الاجتماعية على نحو أفضل.

إن نتيجة التفسير النهائية، هي عادة، تشكيل الآراء التي تُدبّل المعطيات الأساسية، وتُغلق المسار جوهرياً. من الشائع أن تتعارض هذه الآراء بعضها مع بعض، أو أن تكون ضعيفة الإثبات باستعمال المعطيات الأساسية. في مثل هذه الحالات تنحو المجتمعات إلى الدخول في عملية الطرح المُتضارب "التحرير" (يُقدّم مجتمع الموسوعة ويكيبيديا أحد أمثلة

هذه العملية). إننا نحتاج إلى جيل جديد من فضاءات عمل المعطيات يُمكن المساهمين من تذييل المعطيات الأساسية بدليل معطيات تفسيرية، وحل اختلافاتها، وتقنية مجموعات المعطيات، في حين تبقى المعطيات المنفحة جزئياً والمتجانسة جزئياً مُتاحةً للتحري.

تحديات تحليل المعطيات الكبيرة

بعد أن وصّفنا المراحل المُتعددة في مجرى تحليل المعطيات الكبيرة، سننتقل إلى بعض التحديات العامة التي تشمل معظم هذه المراحل، بل كلّها أحياناً، وهذا يعود لخصائص المعطيات الكبيرة. تُظهر المستطيلات الستة في الجزء السفلي من الشكل 2 هذه الخصائص.

التباين (عدم التجانس). عندما يستهلك البشر المعلومات، يمكن بسهولة التسامح في كثير من عدم التجانس. في الحقيقة يوفّر الفرق الدقيق وثراء اللغة الطبيعية عمقاً ذا قيمة. لكن خوارزميات تحليل الآلة تقبل معطيات متجانسة، وهي ضعيفة في إدراك الفروق الدقيقة. ومن ثمّ، تجب هيكلة المعطيات بعناية كخطوة أولى في تحليل المعطيات أو قبل ذلك.

يُعتبر توليد دليل المعطيات الصحيح لوصف المعطيات المُسجّلة على نحو مؤتمت أحد التحديات المُلحقة. فعلى سبيل المثال، قد نحتاج في التجارب العلمية إلى تفاصيل كثيرة تتعلق بالحالات التجريبية الخاصة وبالإجراءات، لنتمكّن من تفسير النتائج على نحو صحيح. تُخفّض نُظْمُ تحصيل دليل المعطيات المسؤولية البشرية في تسجيل الدليل. إذ لا فائدة من تسجيل معلومات عن المعطيات عند تشكّلها إلا إذا أمكن تفسير هذه المعلومات وتوظيفها في مجرى سير تحليل المعطيات. وهذا ما يسمى أصل المعطيات. فمثلاً يمكن لخطأ معالجة في إحدى الخطوات أن يؤدي إلى هدرٍ في تحليل لاحق؛ وبوجود أصل ملائم يمكننا تعرّف كافة المُعالجات اللاحقة المُعتمدة على هذه الخطوة. ومن ثمّ فنحن بحاجة إلى نُظْم معطيات تحافظ على أصل المعطيات ودليل معطياتها خلال إجراءات تحليل المعطيات.

عدم الاكتمال وعدم التوافق. تحتوي المعطيات الكبيرة، على نحو متزايد، معلومات تزوّدها، وعلى نحو متزايد، مصادر متنوعة، ذات موثوقية مُتغيرة. فالغموض والأخطاء والقيم المفقودة مُنتشرة، وتجب معالجتها. في الجانب المُضيء، يمكن استغلال حجم المعطيات الكبيرة والزيادة فيها للتعويض عن المعطيات المفقودة، وللتدقيق التصالبي للحالات المُتعارضة، وإقرار صلاحية العلاقات الموثوقة، وللنظر في التجمعات الكامنة، وكشف العلاقات والنماذج المخفية.

تبرز قضايا مشابهة في استشارات الحشد (crowd-sourcing).⁹ ومع أن كشف وتصحيح معظم الأخطاء يقوم به أشخاص آخرون ضمن الحشد، فإننا نحتاج إلى تقنيات لتسهيل ذلك. بوصفنا بشراً، يمكننا الاطلاع على تقييمات المُنتج، السلبي منها والإيجابي، وأن نخرج بتقييم مجمل يُمكننا، اعتماداً عليه، تقرير شراء هذا المنتج أم لا. نحتاج إلى حواسيب يمكنها أن تُجري المقابلة. تغدو مسائل الغموض والخطأ أكثر وضوحاً في أنماط مُحدّدة من استشارات الحشد المُسماة الاستشعار التشاركي. في هذه الحالة، يمكن لأي شخص معه هاتف نقال أن يكون مُحسناً مُتعدد الأنماط، يجمع أنواعاً مختلفة من المعطيات أنياً (على سبيل المثال، صورة، مقطع فيديو، مقطع صوتي، موقع، سرعة، اتجاه، تسارع). يكمن التحدي الإضافي هنا في عدم التيقن بأجهزة جمع المعطيات. يُمكن استغلال كون المعطيات المجموعة مترابطةً مكانياً

⁹ استشارات الحشد (متعدد المصادر): مُصطلح حديث في مجال الأعمال أُدخل عام 2005. ويُعرّف بأنه إجرائية الحصول على الخدمات أو الأفكار أو المضامين المطلوبة بمساهمات استشارية من مجموعة كبيرة من الأشخاص، وعلى وجه الخصوص من مُجتمع الوِب. (المرجم)

وزمنياً لتقييم صحتها على نحو أفضل. عندما تُجمع المعطيات بهدف الاستئجار كما في حالة "التركي الميكانيكي"¹⁰ فإن دوافع العاملين المتنوعة تؤدي إلى ظهور نموذج آخر من الخطأ.

من المرجح، بعد تطبيق إجراءات تصحيح الأخطاء، بقاء بعض الأخطاء وعدم الاكتمال. ولا بد من معالجة هذه الأخطاء وعدم الاكتمال إبان تحليل المعطيات. إن فعل هذا على نحو صحيح يعتبر تحدياً. ثمة طريقة لتحقيق تقدّم في هذا المجال يعرضها عمل حديث عن معالجة المعطيات الاحتمالية المتضاربة والاستعلام عنها.

التقييس (scale). إن أول ما يخطر ببال أحدنا عن المعطيات الكبيرة هو حجمها بطبيعة الحال. ما تزال معالجة حجوم كبيرة وتزداد باطراد، لمعطيات، تشكل تحدياً على مدى عدة عقود. كان هذا التحدي يُخفّف بزيادة سرعة المعالجات، اعتماداً على قانون مور. لكن ثمة تغيير أساسي يحدث الآن: فحجم المعطيات يزداد بسرعة أكبر من زيادة سرعات المعالج وموارد الحاسوب الأخرى.

وبسبب معوقات الطاقة، فقد تباطأت الساعات، وتُنبت المعالجات بعدد متزايد من النوى. باختصار، علينا التعامل بازدواجية ضمن العقدة الواحدة. إلا أن تقنيات المعالجة المتوازية للمعطيات التي اعتمدت في السابق لمعالجة المعطيات في العقد، لا تُطبّق - ويا للأسف - مباشرة على الازدواجية ضمن العقدة، حيث البنيان مختلف جذرياً. فعلى سبيل المثال، هناك عدة موارد لعنديات من قبيل خواصي المعالج وقنوات ذاكرة المعالج يجري تشاركها ضمن النوى في ذات العقدة.

ثمة تحوّل مثير آخر يتمثل في التحوّل إلى الحوسبة السحابية، التي تجمع مهام أعمال متعددة ذات أهداف أداء متنوعة ضمن حشد كبير جداً. يُجهد هذا المستوى من تشارك الموارد في حشود غالبية وكبيرة تقنيات حوسبة الشبكات والحشد القديمة، وهو يحتاج إلى طرق جديدة لتحديد كيفية تشغيل وتنفيذ أعمال معالجة المعطيات بحيث نحقق الأهداف لكل مهمة على نحو غير مُكلف، وبحيث نتعامل مع إخفاقات النظام التي يتكرر حدوثها على نحو أكبر كلما شغلنا على نُظم أكبر وأكبر.

يقود هذا إلى الحاجة لاستمثال شامل لبرامج المستخدمين المتعددة، تشمل تلك التي تؤدي مهام تعلم آلة معقدة. غالباً ما يؤدي الاعتماد على استمثال المُستخدم للبرامج إلى استثمار ضعيف للحشد، إذ إن المستخدمين يجهلون برامج المستخدمين الآخرين بسبب الافتراضية. يحتاج الاستمثال الشامل في النظام إلى أن تكون البرامج شفافة بقدر كافٍ، كما في أنظمة قواعد المعطيات الارتباطية، على سبيل المثال، حيث تُصمّم لغات الاستعلام التفسيرية مع الأخذ بالحسبان هذا الأمر. في الحقيقة، إذا بنى المستخدمون ممرات معالجة معقدة للمعطيات الكبيرة، فمن الضروري أن يكون لديهم معلومات بدائية عالية المستوى مناسبة لتوصيف احتياجاتهم.

إضافة إلى الأسباب التقنية لمزيد من التطوير لمقاربات تفسيرية لتحليل المعطيات الكبيرة، فإن هناك ضرورة تجارية كذلك. فغالباً ما تُعهد المنظمات معالجة المعطيات الكبيرة أو جوانب متعددة منها لجهات خارجية. وللوصول إلى اتفاقات ذات معنى وقابلة للتنفيذ، نحتاج إلى مواصفات تفسيرية، إذ إن الأمر الهام في توكيل التنفيذ لجهة خارجية، هو أن توصّف بدقة أية مهمة ستؤدي دون الدخول في تفاصيل كيفية التنفيذ.

¹⁰ Mechanical Turks: سوق على الإنترنت يعتمد مبدأ المساهمات الاستشارية (الحشد) ويسمح للأفراد والشركات بتسويق استعمال الذكاء البشري لأداء مهام لا تستطيع الحواسيب إجراؤها في الوقت الراهن. ويعود أصل التعبير "Turk" إلى آلة مزيّفة للعبة الشطرنج ابتكرت في القرن الثامن عشر، ويُشغلها لاعب شطرنج مُحترف مُختبئ بداخلها. (المترجم)

التوقيت المناسب (Timeliness). مع ازدياد حجم المعطيات، نحتاج إلى تقنيات في الزمن الحقيقي لتجزئ وتُصَفِّي ما يجب تخزينه، إذ إنه من غير الممكن اقتصادياً تخزين المعطيات الخام. يظهر هنا تحدّي مُعدّل التحصيل الذي جرى وصفه سابقاً، وتحدّي التوقيت المناسب الذي سنناقشه فيما يأتي. فعلى سبيل المثال، إذا اشْتَبِه بعملية احتيال مالية جرت من بطاقة ائتمان، فمن المفترض، في الحالة المثالية، تعرّفها قبل اكتمال العملية- بل ربما منع حدوث العملية كلياً. من الواضح أن تحليلاً كاملاً لتاريخ مشتريات المستخدم لن يكون ممكناً في الزمن الحقيقي. إننا نحتاج إلى نتائج قبلية، بحيث يمكن باستعمال قليل من حسابات تزايدية بمعطيات جديدة أن نصل إلى تحديد سريع. يتمثل التحدي الرئيسي في توفير أزمنة استجابة تفاعلية للاستعلامات المُعدّدة بما يتناسب مع تدفقات الحوادث بحجوم كبيرة.

إن إيجاد عناصر تُطابق معياراً محدداً ضمن مجموعة معطيات كبيرة، يُشكّل نمطاً شائعاً آخر. يحدث هذا النوع من البحث على نحو متكرر في سياق تحليل المعطيات. من الواضح أن مسح كامل مجموعة المعطيات لإيجاد عناصر مناسبة غير عملي على الإطلاق. وأفضل من ذلك إنشاء بنى استدلالية سلفاً لإيجاد العناصر المؤهّلة بسرعة. لنفرض مثلاً نظام إدارة مرور يتضمّن معلومات عن آلاف المركبات وعن النقاط المزدحمة على الطرق. قد يحتاج النظام إلى أن يتنبأ بنقاط الازدحام المُحتملة على طريق اختاره أحد المستخدمين، وأن يقترح البدائل. يتطلّب هذا الأمر تقييماً لاستعلامات متعددة مجاورة مكانياً بالتعامل مع مسارات عناصر متحركة. إننا نحتاج إلى ابتكار بنى استدلالية للتعامل مع تنوع واسع من هذه المعايير.

الخصوصية وملكية المعطيات. تشكل خصوصية المعطيات أمراً آخر بالغ الأهمية، وتزداد أهميتها في إطار المعطيات الكبيرة. ففي سجلات الصحة الإلكترونية ثمة قوانين صارمة تحدد المعطيات التي يمكن كشفها في السياقات المختلفة. أما في حالة المعطيات الأخرى، فإن القوانين، في الولايات المتحدة على وجه الخصوص، هي أقل صرامة. ومع ذلك، فثمة خوف عام كبير يتعلق بالاستعمال غير الملائم للمعطيات الشخصية، خصوصاً من خلال ربط المعطيات من مصادر متعددة. إن إدارة الخصوصية على نحو فعّال مشكلة فنية واجتماعية على حد سواء، ومن ثمّ تجب معالجتها من وجهتي النظر معاً لتحقيق وعدّ المعطيات الكبيرة.

لنأخذ، على سبيل المثال، معطيات جُمعت من خدمات تعتمد على الموقع، وتتطلب من المستخدم مشاركة موضعه مع مزود الخدمة. ثمة مخاوف جلية تتعلق بالخصوصية، لا يمكن معالجتها بإخفاء هوية المستخدم فقط دون إخفاء موضعه. إذ يمكن لمهاجم أو لوجود مخدّم خبيث يعتمد على الموقع أن يستخلص هوية مصدر الاستعلام من معلومات موضعه المُلحقة. فمثلاً، من المحتمل أن يترك المستخدم أثراً يمكن ربطه بموضع سكن معيّن أو مكتب، يُستعمل لتحديد هوية المستخدم. ثمة أنواع عديدة من المعلومات الخاصة من قبيل قضايا الصحة (وجود الشخص في مركز معالجة السرطان على سبيل المثال) أو قضايا دينية (مثل الحضور في كنيسة) يمكن كشفها فقط بمراقبة حركة شخص ما ونمط تكرارها بمرور الزمن. وبوجه عام، يوجد ترابط قوي بين هويات الأشخاص وأنماط تحركاتهم.¹¹ ولكن في الخدمات التي تعتمد الموقع، نحتاج إلى معرفة موضع المستخدم للولوج الناجح إلى المعطيات ولجمعها، وبناءً عليه، فإن القيام بذلك على نحو صحيح يعتبر تحدياً.

ثمة قضية أخرى وهي أن كثيراً من الخدمات المباشرة اليوم تتطلّب تشارك معلومات خاصة (مثل تطبيقات الفيسبوك)، لكننا، بعيداً عن التحكم بمستوى النفاذ إلى السجلات، لا ندرك ماذا يعني تشارك المعطيات، وكيف يمكن ربط المعطيات المتشاركة، وكيف نوفر للمستخدمين تحكماً ناعماً بهذا التشارك بطريقة بديهية ولكنها فعالة. إضافة إلى ذلك،

المعطيات الحقيقية ليست ساكنة لكنها تكبر وتتغير بمرور الزمن، ولم تنتشر أية نتائج صادرة عن التقنيات الشائعة، في أي محتوى مفيد، ضمن هذا السيناريو.

تعتبر الخصوصية واحداً من جوانب ملكية المعطيات. وعموماً، مع ازدياد تقدير قيمة المعطيات، أصبحت قيمة المعطيات التي تملكها هيئة، تحظى باهتمام استراتيجي مركزي. تهتم الهيئات بكيفية زيادة هذه المعطيات، مع المحافظة على المزايا الفريدة لها. كما غدت أسئلة من قبيل كيف تُشارك أو تبيع المعطيات دون فقد التحكم هامة. لا تختلف هذه الأسئلة كثيراً عن قضايا إدارة الحقوق الرقمية (Digital Rights Management: DRM) التي تواجهها صناعة الموسيقى حيث تحوّل التوزيع من بيع وسائط مادية مثل الأقراص المُدمجة إلى الشراء الرقمي؛ وهكذا فنحن بحاجة إلى مقاربات فعالة ومرنة في إدارة الحقوق الرقمية.

المنظور البشري: التصور (visualization) والتعاون. لكي تصل المعطيات الكبيرة إلى كامل إمكاناتها، علينا أن نأخذ في الحسبان ليس فقط متطلبات النظام بل رؤى الأشخاص أيضاً. علينا أن نتحقق أنّ النقاط النهائية - البشر - يمكنها أن "تستوعب" نتائج التحليل على نحو مناسب، وأن لا تضيق هذه النتائج في بحر المعطيات. فعلى سبيل المثال، تساعد خوارزميات التوبيخ والتزكية في تعرّف المعطيات الأكثر إثارة لاهتمام المستخدم، آخذة بالاعتبار أفضليّاته. ولكن عند استعمال هذه التقنيات للاكتشاف والاستقصاء العلمي، يجب الانتباه إلى حصر المستخدمين ضمن "فقاعة الترشيح"^{11,21} لمعطيات مشابهة لما رآه في السابق - فقد جاء الكثير من الاكتشافات المثيرة للاهتمام من كشف أشخاص لا يسكنون في مكان عملهم.

على الرغم من النقص الهائل في التحليل الحاسوبي، ثمة أنماط كثيرة يمكن للبشر كشفها بسهولة، على حين يصعب على خوارزميات الحاسوب إيجادها. فعلى سبيل المثال، يستثمر البرنامج CAPCHA¹² تماماً هذه الحقيقة للتمييز بين مستخدم الوب، وهل هم من البشر أم من البرامج الحاسوبية. لن تكون تحليلات المعطيات الكبيرة جميعها حاسوبية - بل سيجري تصميمها لتشمل العامل البشري في دورتها. يحاول الحقل الفرعي الخاص بالتحليل المرئي القيام بذلك، على الأقل بما يتعلق بمرحلة النمذجة والتحليل في المسار. وهناك قيمة مشابهة لإدخالات الإنسان في كافة مراحل مسار التحليل.

في العالم المعقّد اليوم، كثيراً ما نحتاج إلى خبراء متعددين من مجالات مختلفة لإدراك حقيقة ما يجري. على نظام تحليل المعطيات الكبيرة أن يقبل إدخالات من خبراء أدميين متعددين، وأن يتشارك في تحزّي النتائج. يمكن فصل هؤلاء الخبراء المتعددين مكانياً وزمنياً عندما يكون مكلفاً جمع الفريق كاملاً في مكان واحد. ويجب أن يقبل نظام المعطيات دخل الخبراء الموزّع، ويدعم تعاونهم. يحتاج ذلك، تقنياً، إلى أن نتشارك فيما هو أكثر من المعطيات الخام فقط؛ إذ علينا أيضاً أن نبحث في كيفية تمكين تشارك خوارزميات وأدوات من قبيل نتائج تجريبية (على سبيل المثال، تطبيق خوارزمية ذات قيم وسائط محددة على لقطة مُعطاة من مجموعة معطيات متجددة).

¹¹ filter bubble: هي نتيجة يحصل عليها المُستعمل من الإنترنت عندما تقوم خوارزميات الموقع بتخمين ما يرغب أن يراه المُستعمل وذلك اعتماداً على معلومات عن هذا المستخدم (المستعمل) من قبيل (موضعه، واستعلاماته السابقة، وتاريخ بحثه...)، ومن ثم يصبح المستخدمون مفصولين عن المعلومات التي تخالف وجهة نظرهم، أي يجري عزلهم ضمن محيطهم الثقافي أو (فقايعهم العقائدية). (المترجم)

¹² برنامج يهدف إلى تمييز إدخالات البشر عن إدخالات الآلة، ويهدف إلى منع استخلاص معلومات من مواقع الوب على نحو مؤتمت. والاسم الكامل لهذا البرنامج " اختبار تورنغ العام المؤتمت كاملاً لتمييز البشر من الحواسيب " وهو اختبار لاستجابة المستخدم لأسئلة عن أشكال وأرقام يحدد في ضوئها إذا كان المستخدم بشراً أم برنامجاً حاسوبياً. (المترجم)

أصبحت النظم التي تضم مؤثرات بصرية ذات لوحة ألوان غنية يمكن إنشاؤها بسرعة ووضوح، مهمة في نقل نتائج الاستعلام إلى المستخدمين بطرق تُفهم جيداً في الحقل الخاص، وتحمل تفاصيل مناسبة. وعلى حين كان مستخدمو نظم الاستعلام الذكي سابقاً راضين بعروض مُجدولة، يرغب المحللون في هذه الأيام في جمع النتائج وعرضها بمؤثرات بصرية قوية تساعد على التوضيح، وتدعم مشاركة المستخدم. إضافة لذلك، يستطيع المستخدم بنقرات محدودة أن يتوغل عميقاً في كل قطعة من المعطيات ليراها ويفهم أصلها. وهذا مهم خاصة مع وجود عدد متزايد من الأشخاص لديهم معطيات يرغبون في تحليلها.

تعاونيات المعطيات الكبيرة. مع ازدياد المجتمعات التي تعتمد على إدارة المعطيات باستعمال حوسبة السحابة، ومع تحول الكثير من مخازن المعطيات إلى موارد مفتاحية، فإن القيمة المتوقعة للتعاون باستعمال المعطيات المُتشاركة تزداد باطراد. والسؤال هو كيف نسمح للمستخدمين بإنشاء تحليل معطيات يجمع بين معطياتهم والمعطيات المُتشاركة، والسماح لمستخدمين آخرين (انتقائياً) بتشغيل وتعديل وإعادة توزيع أدوات التحليل هذه والتي يمكن أن تتفاوت من استفسارات وحيدة ولغاية إنجاز مخططات عمل نمذجة واستخلاص نتائج متكاملة؟ يتطلب ذلك معالجة عدد من المسائل (مثل المنشأ، والتحكم بالنفاد، وانسياب العمل) لكنه يحمل إمكانية كبيرة لتعاون متزايد، ويرفع مستوى الشفافية في العمل التعاوني (تصور أن تتمكن من إعادة تنفيذ كافة التحاليل المذكورة في مقالة مُستعملاً ذات المعطيات والرّماز اللذين استعملهما مؤلفوها، وأن تتمكن من تهذيب النتائج وإعادة نشرها!).

ثمة طريقة جديدة شائعة لتسخير المهارة البشرية لحل المشاكل تجري عن طريق ترويج المساهمات. تمثل الموسوعة ويكيبيديا الموصولة مباشرة على الإنترنت أفضل مثال معروف للمعطيات المُروّجة بهذه الطريقة. تحمل المقاربات الاجتماعية لتحليل المعطيات الكبيرة وعوداً عظيمة. فعندما نجعل طيفاً واسعاً من الأدوات المتمركزة حول المعطيات قابلاً للتشارك، فإننا نفتح الباب للآليات الاجتماعية التي من قبيل تبويب الأدوات، ولوحات القيادة (على سبيل المثال، مقارنة فعالية خوارزميات متعددة على مجموعة المعطيات ذاتها)، والسُّمعة المُستحثة للخوارزميات والخبراء.

الخلاصة

لقد دخلنا حقبة المعطيات الكبيرة. وكثير من قطاعات اقتصادنا تتجه الآن إلى نموذج اتخاذ القرار المعتمد على المعطيات حيث تعتمد نواة الأعمال على تحليل حجوم كبيرة ومتنوعة لمعطيات تُنتج باستمرار. لدى هذا العالم المُعتمد على المعطيات القدرة على تحسين مردود الشركات، وتحسين نوعية حياتنا. ولكن، ثمة تحديات لا بد من التصدي لها لنتمكن من استثمار الإمكانيات الكاملة للمعطيات الكبيرة. سلط هذا المقال الضوء على التحديات التقنية المفتاحية الواجب التصدي لها، وبيّن وجود تحديات أخرى، اقتصادية واجتماعية وسياسية، لم تجر معالجتها في هذا المقال لكن لا بد من التصدي لها أيضاً. لا تظهر كافة التحديات التقنية التي طُرحت هنا في كافة سيناريوهات التطبيق. إلا أن كثيراً منها يظهر. كما أن حلول التحدي قد لا تكون ذاتها في كافة الحالات. لكن عموماً ثمة تشابهات كافية تدعم التعلّم بالمقارنة. كما أن الطيف العريض للتحديات المعروضة هنا تُشكّل عناوين جيدة للبحث في مجالات عدة من علم الحاسوب. لقد جمعنا بعض الاقتراحات

للتوسع بالقراءة في الموقع: <http://db.pitt.edu/bigdata/resources>. وهو يضم عشرات المقالات اختيرت من حيث الشمولية والأهمية، بدلا من دراسة مرجعية شاملة، ربما تتألف من آلاف المقالات.

شكر خاص

اعتمد هذا المقال على مقالة أصلية¹³ كتبها كثير من الباحثين البارزين الذين تُقدّر مساهماتهم. شكراً لكل من: Divyakant Agrawal، و Philip Bernstein، و Elisa Bertino، و Susan Davidson، و Umeshwar Dayal، و Michael Franklin، و Laura Haas، و Alon Halevy، و Sam Madden، و Kenneth Ross، و Dan Suci، و Shiv Vaithyanathan، و Jennifer Widom.

جرى تمويل عمل البروفسور H.V.Jagdish من مَنح المؤسسة NSF: 13 IIS 1017296، و OIA-1028162، و IIS 1250880. وجرى تمويل عمل البروفسور Alexandros Labrinidis من مَنح مؤسسة العلوم الوطنية: IIS-0746696، و IIS 1017149، و CBET-1250171. وجرى تمويل البروفسور Yannis Papakonstantinou من مَنح مؤسسة العلوم الوطنية: IIS-1117527، و SHB- 1237174، و DC-0910820، ومن جائزة أبحاث المعلوماتية. وجرى تمويل عمل البروفسور Jignesh M. Patel من مَنح مؤسسة العلوم الوطنية: III-0963993، و IIS-1250886، و IIS-1110948، و CNS-1218432، ومن هبات من غوغل، ومن الشركة Johnson Controls ومن الشركة Microsoft، ومن الشركة Symantec، ومن الشركة Oracle، وجرى تمويل عمل البروفسور Cyrus Shahabil من منحة مؤسسة العلوم الوطنية IIS-1115153، وعقد مع شركة مترو لوس انجلوس، ومنحة مالية مفتوحة من الشركة Microsoft ومن الشركة Oracle.

كافة الأفكار، أو الاستنتاجات، أو النتائج، أو التوصيات المُعبّر عنها في هذا المقال هي مسؤولية كاتبها.

المراجع

- [1] Computing Community Consortium. Advancing Discovery in Science and Engineering. Spring 2011.
- [2] Computing Community Consortium. Advancing Personalized Education. Spring 2011.
- [3] Computing Community Consortium. Smart Health and Wellbeing. Spring 2011.
- [4] Computing Community Consortium. A Sustainable Future. Summer 2011.
- [5] Computer Research Association. Challenges and Opportunities with Big Data. Community white paper available at <http://cra.org/ccc/docs/init/bigdatawhitepaper.pdf>
- [6] Dobbie, W. and Fryer, Jr. R.G. Getting Beneath the Veil of Effective Schools: Evidence from New York City. NBER Working Paper No. 17632. Issued Dec. 2011.
- [7] *Economist*. Drowning in numbers: Digital data will flood the planet—and help us understand it better. (Nov 18, 2011); <http://www.economist.com/blogs/dailychart/2011/11/big-data-0>
- [8] Flood, M., Jagdish, H.V., Kyle, A., Olken, F. and Raschid, L. Using data for systemic financial risk management. In Proc. 5th Biennial Conf. Innovative Data Systems Research (Jan. 2011).

¹³ NSF: National Science Foundation، مؤسسة العلوم الوطنية: وكالة حكومية أمريكية تدعم البحوث الأساسية والتعليم في كافة مجالات العلوم والهندسة فيما عدا البحوث الطبية. وهي تدعم ما يقارب 20% من الأبحاث في الولايات المتحدة. (المترجم)

- [9] Forbes. Data-driven: Improving business and society through data. (Feb. 10, 2012); <http://www.forbes.com/special-report/data-driven.html>
- [10] Gartner Group. Pattern-Based Strategy: Getting Value from Big Data. (July 2011 press release); <http://www.gartner.com/it/page.jsp?id=1731916>
- [11] González, M.C., Hidalgo, C.A. and Barabási, A-L. Understanding individual human mobility patterns. *Nature* 453, (June 5, 2008), 779–782.
- [12] Hey, T., Tansley, S. and Tolle, K., eds. *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft Research, 2009.
- [13] Kahn, S.D. On the future of genomic data. *Science* 331, 6018 (Feb. 11, 2011), 728–729.
- [14] Lazar, D. et al. Computational social science. *Science* 323, 5915 (Feb. 6, 2009), 721–723.
- [15] Lohr, A. The age of Big Data. *New York Times* (Feb. 11, 2012); <http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html>
- [16] Lohr, S. How Big Data became so big. *New York Times* (Aug. 11, 2012); <http://www.nytimes.com/2012/08/12/business/how-big-data-became-so-big-unboxed.html>
- [17] Manyika, J. et al. *Big Data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute. May 2011.
- [18] National Science and Technology Council. *Materials Genome Initiative for Global Competitiveness*. June 2011.
- [19] Noguchi, Y. Following the Breadcrumbs to Big Data Gold. *National Public Radio* (Nov. 29, 2011); <http://www.npr.org/2011/11/29/142521910/the-digital-breadcrumbs-that-lead-to-big-data>
- [20] Noguchi, Y. The Search for Analysts to Make Sense of Big Data. *National Public Radio*, (Nov. 30, 2011); <http://www.npr.org/2011/11/30/142893065/the-search-for-analysts-to-make-sense-of-big-data>
- [21] Pariser, E. *The Filter Bubble: What the Internet Is Hiding From You*. Penguin Press, May 2011.
- [22] PCAST Report. *Designing a Digital Future: Federally Funded Research and Development in Networking and Information Technology* (Dec. 2010); <http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-nitrd-report-2010.pdf>
- [23] SDSS-III: *Massive Spectroscopic Surveys of the Distant Universe, the Milky Way Galaxy, and Extra-Solar Planetary Systems* (Jan. 2008); <http://www.sdss3.org/collaboration/description.pdf>

المؤلفون

- H.J. Jagadish** (jag@umich.edu). أستاذ جامعي للهندسة الكهربائية وعلم الحاسوب في جامعة ميشيغان في أن آر بور.
- Johannes Gehrke** (johannes@cs.cornell.edu). أستاذ قسم علم الحاسوب في جامعة كورنيل في نيويورك.
- Alexandros Labrinidis** (labrinids@cs.ucsd.edu). أستاذ مساعد في قسم علم الحاسوب في جامعة بيتسبرغ ومساعد مدير مخبر تقنيات إدارة المعطيات المتقدمة.
- Yannis Papakonstantinou** (yannis@cs.ucsd.edu). أستاذ علم الحاسوب والهندسة في جامعة كاليفورنيا في سنت ياغو.
- Jignesh M. Patel** (jagnesh@cs.ucsd.edu). أستاذ علم الحاسوب في جامعة وينسكونسن في ماديسون.
- Raghu Ramakrishnan** (raghu@microsoft.com). زميل تقني والمدير الفني لخدمات المعلومات في شركة مايكروسوفت بمدينة ريدموند.
- Cyrus Shahabi** (shahabi@usc.edu). أستاذ علم الحاسوب والهندسة الكهربائية ورئيس مخبر المعلومات في جامعة كاليفورنيا الجنوبية، ومدير مركز نظم الإعلام المتكاملة التابع لمنظمة العلوم الوطنية في أمريكا.

الأمن والخصوصية لأنظمة الحقيقة المزيدة

SECURITY AND PRIVACY FOR AUGMENTED REALITY SYSTEMS*

Franziska Roesner, Tadayoshi Kohno, David Molnar

ترجمة: د. خالد مصري
مراجعة: د. عمران قوبا

تطرح أنظمة الحقيقة المزيدة مخاوف أمنية محتملة يجب التصدي لها قبل أن تصبح هذه الأنظمة واسعة الانتشار.

تعد تقانات الحقيقة المزيدة (AR) augmented reality بتعزيز إدراكنا للعالم الحقيقي وتفاعلنا معه. وعلى خلاف أنظمة الحقيقة الافتراضية، التي تستبدل عالماً محاكى بالعالم الحقيقي، تتحس أنظمة الحقيقة المزيدة خواص العالم المادي وتضع الإشارات المرئية والصوتية والمسبية التي يولدها الحاسوب فوق تغذية راجعة من العالم الحقيقي في الزمن الحقيقي. سنعرض في هذه المقالة، المخاوف المتعلقة بالأمن والخصوصية المرتبطة بأنظمة الحقيقة المزيدة نفسها، إضافة إلى تلك التي تنشأ عن التقانات الداعمة.

استكشف الباحثون فكرة الحقيقة المزيدة منذ الستينيات، عندما وصف إيفان سوذرلاند Ivan Sutherland جهاز رؤية شفافاً مثنياً على الرأس يُظهر معلومات ثلاثية الأبعاد¹⁰. منذ التسعينيات، ركزت الحقيقة المزيدة، لكونها مجال بحث، على التغلب على تحديات تقانة الإظهار، والتعقب، والتسجيل لصَف الأشياء الحقيقية والافتراضية على نحو صحيح، وتحديات واجهات المستخدم والعوامل البشرية، وتجهيزات التحسس المساندة، وتصميم تطبيقات جديدة للحقيقة المزيدة. 413622و620364
مع ذلك، لم تُطرح تقانات الجيل الأول لأنظمة الحقيقة المزيدة تجارياً، إلا في الآونة الأخيرة. على سبيل المثال، أصدرت شركة غوغل Google حديثاً عدداً محدوداً من نظارات غوغل Google Glass الرأسية heads-up لتطبيقات الحقيقة المزيدة. يوجد العديد من التطبيقات الأخرى للجيل الأول لأنظمة الحقيقة المزيدة، أتاحها انتشار الهواتف الذكية والتجهيزات النقالة الأخرى. من هذه الأمثلة، تطبيق الآيفون Word Lens iPhone، وهو تطبيق يضع النص المترجم فوق النص الأجنبي الواقع في حقل رؤية الكامرة (آلة التصوير)، والتطبيق Layar، وهو منصة حقيقة مزيدة تعتمد على تحديد الموقع الجغرافي بواسطة الإنترنت geolocation تُمكن المطورين من توليد طبقات الحقيقة المزيدة للعالم (مثلاً، للعب

* نُشر هذا البحث في مجلة Communications of the ACM، المجلد 57، العدد 4، نيسان (أبريل) 2014، الصفحات 88 – 96.



الشكل 1. الحقيقة المزيدة المعتمدة على الهاتف. على اليسار، صورة
World Lens، وهو تطبيق للهاتف النقال "الأيفون" يوفر ترجمة سلسلة "في
الصورة" in-picture (المصدر: [http://www.flickr.com/photos/
neven/5269418871/](http://www.flickr.com/photos/neven/5269418871/)). يظهر في الصورة ترجمة التطبيق لكلمة "craft"
من الإنكليزية إلى الإسبانية، ثم العودة مرة أخرى. على اليمين، صورة
Layar، وهو "متصفح حقيقة مزيدة" تُرَوِّد به هواتف أندرويد
(المصدر: [http://site.layar.com/company/blog/make-your-
own-layar-screen-shot-with-the-dreamcatcher](http://site.layar.com/company/blog/make-your-own-layar-screen-shot-with-the-dreamcatcher)).

كمرة فيديو شغالة دائماً توضع فوق الأذن - سماتٍ تُمكن حاملها من مشاركة ما يشاهدون من مشاهد فيديو مباشرة مع
أي شخصٍ آخر في العالم. كما أتاح طقم التطوير (SDK) Microsoft Software Development Kit لجهاز الكينكت



الشكل 2. دخل وخرج قابلان للارتداء. على اليسار، كمرة Looxcie يمكن
ارتداؤها، يرتديها حارس غابات في كينيا (المصدر: <http://Looxcie.com/index.php/image-gallery>).
على اليمين، نموذج أولي لنظارات
غوغل في حزيران 2012 (المصدر: [http://www.flickr.com/photos/
azugaldia/7457645618/](http://www.flickr.com/photos/azugaldia/7457645618/)).

(game playing)، انظر الشكل 1. إن ظهور المعالجات 1 GHz ومُجسَّات
الموضع والكميرات ذات الضبط التلقائي والميز العالي في الهواتف النقال، في
الأونة الأخيرة، جعلَ هذه التطبيقات
ممكنة.

سنلقي في هذه المقالة نظرةً على
فضاء الحقيقة المزيدة، آخذين بالحسبان
كلاً من التطبيقات المباشرة لأنظمة
الحقيقة المزيدة والتقانات اللازمة لدعم
هذه التطبيقات. تتوفر حالياً تجهيزات،
غير الهواتف النقال، تُعزِّز التحسس
والإظهار ومشاركة المعطيات، مما
سيُتيح أنظمة للحقيقة المزيدة أكثر
تعقيداً. فمثلاً، يتضمن Looxcie - وهو

kinect²⁰ - الذي يوفر تحسناً دقيقاً
للحركة بدمج كمرة RGB وكمرة عمق
depth وصفيفة متعددة multi-array
من المكروفونات - العديد من النماذج
الأولية لتطبيقات الحقيقة المزيدة. إضافةً
إلى نظارات غوغل، تُوفر حالياً عدة
شركات أجهزة إظهار شفافة ولبوسة
wearable لأغراض بحثية، من هذه
الشركات شركة Lumus Vuzix و
Meta SpaceGlasses. يبين الشكل
2 أمثلةً على تجهيزات الدخل والخرج
هذه. (يعرض الجدول 1 ملخصاً لتقانات

¹ جهاز يُربط بالجهاز الأم يحتوي على كمرتين ومكروفونات. تلتقط الكمرة الأولى صوراً ملونة، في حين تستعمل الثانية تقانة مُجس العنق ثلاثي
الأبعاد 3D depth sensor وهي التي تمكن من التقاط العمق. يبعث جهاز Kinect أشعة تحت حمراء، وتُمكن كمرة الأشعة تحت الحمراء رؤية
كل المشهد. يُمكن الجمع بين كل هذا جهاز Kinect من التقاط أصوات الشخص الحاضر في الغرفة ومكانه. (المترجم).

تمكين الحقيقة المزيدة؛ يُباع العديد من هذه التقانات حالياً، في حين ما يزال بعضها في المرحلة التجريبية.

الجدول 1. ملخص تقانات الحقيقة المزيدة البازغة والتجارية.

التقانات التجريبية فقط	التقانات المتوفرة تجارياً حالياً	
مُحسَّات لمسية ²⁹	كميرات RGB قابلة للارتداء GPS (خطأ 5 أمتار أو أكثر) تحسس حركة دقيق (مثلاً، Kinect)	المُحسَّات (الدخل)
مظهار شفاف قريب من العين أجهزة إظهار مُتضمنة (عدسات لاصقة ¹²) تغذية راجعة لمسية ²⁷	مظهار مُعتم قريب من العين مظهار/ مكبر صوت للهاتف سماعة أذن bluetooth غير مرئية	التغذية الراجعة (الخرج)
خدمات سحابة معقدة (تعرف الأغراض) تعقب دون علامة تعرف وجه جيد نسخ transcription رخيص ودقيق	خدمات سحابة cloud بسيطة (معرض صور) تعقب معتمد على العلامة ¹⁹ Markerbased كشف جيد عن الوجه (ليس تعرفاً) ¹⁷ نسخ transcription غالي أو رخيص الثمن ولكن غير دقيق	الخدمات
مشاركة آلية	مشاركة انتقائية (صور وفيديو ومواقع)	المشاركة

سُتْمَكُن هذه التقانات من ظهور تطبيقات تجارية لأنظمة الحقيقة المزيدة، وهي على أعتاب ابتكار كبير سيجلب منافع كبيرة لكثير من المستخدمين. مع ذلك، يمكن لهذه التقانات أن تجلب أيضاً مخاطر حاسوبية غير متوقعة للأمن والخصوصية. لم تتناول الأبحاث السابقة في أنظمة الحقيقة المزيدة هذه القضايا إلا فيما ندر. وبدلاً من انتظار نضوج هذه التقانات نضوجاً تاماً، ومحاولة تطوير حمايات الأمن والخصوصية بأثر رجعي، فإننا نرى أن الآن هو الوقت المناسب لتناول قضايا الأمن والخصوصية، حيث لا تزال التقانات فتيةً ومطواعةً. لتوجيه هذه العملية، نطرح الأسئلة الآتية: ماهي التحديات البحثية الجديدة لقضايا الأمن والخصوصية التي تفرزها أنظمة الحقيقة المزيدة، وما هي التقانات التي تدعمها؟ ماهي الفرص الجديدة التي تولدها تقانات الحقيقة المزيدة لتحسين الأمن والخصوصية؟

نكتشف أن تقانات الحقيقة المزيدة تُشكّل أرضاً جديدةً خصبةً وهامةً لأبحاث الحاسوب وصناعة أمنه وخصوصيته. بالطبع، يجب أن تُعزَّز هذه التقانات الممارسات الفضلى القياسية للأمن، مثل التعمية على مستوى الجهاز والشبكة. ومع ذلك، نجد عقبات فريدة من نوعها - مثل معالجة التضارب بين التطبيقات المتعددة التي تتشارك نفس خرج نظام الحقيقة المزيدة - وهي تُمثّل في الوقت نفسه تحدياً فكرياً، يمكن التغلب عليها. توجد تحديات أخرى، مثل التحكم في نفاذ المعطيات، وهي معروفة جيداً في مجالات أخرى غير أنها تصبح أكبر أهميةً في تقانات الحقيقة المزيدة لأن مداخلها في حالة تشغيل وتحسس دائمين. ونظراً للأهمية المستقبلية لتقانات الحقيقة المزيدة، يمكن للباحثين الذين تصدوا لهذه القضايا سابقاً في مجالات أخرى أن يجدوا فائدةً في إعادة تركيز اهتمامهم على تطبيقات الحقيقة المزيدة.

إن أنظمة الحقيقة المزيدة إضافةً إلى أنها تطرح تحدياتٍ جديدة، تطرح فرصاً لتطبيقاتٍ جديدة تُحسِّن الأمن والخصوصية. مثلاً، يمكن لهذه التقانات أن توفر رؤىً رقميةً شخصيةً للمحتويات على أجهزة الإظهار الشخصية. لنتخيل مديراً لكلمات المرور يضع مؤشرات مرئية فوق المفاتيح الصحيحة لكلمة مرور معقدة عندما ينظر المستخدم إلى لوحة المفاتيح، أو لنتخيل تطبيقاً يُنبه المستخدم عندما يكذب شخصٌ ما.

سنستكشف في هذه المقالة التحديات الجديدة للأمن والخصوصية التي تولدها تقانات الحقيقة المزيدة، والتوجهات الدفاعية، والتطبيقات الجديدة لأنظمة الحقيقة المزيدة لقضايا الأمن والخصوصية المعروفة.

التحديات

تتميز تطبيقات الحقيقة المزيدة وتقاناتها التي سندرسها بإحدى الخصائص الآتية، أو ببعضها، إضافةً إلى التعريف التقليدي لمحاذاة الأشياء الحقيقية والافتراضية في الزمن الحقيقي:

- مجموعة معقدة من تجهيزات الدخل والمُحسَّات التي تكون في حالة تشغيل دائم (مثلاً، كَمِرَّة ونظام تحديد المواقع العالمي (GPS) Global Positioning System ومكروفون).
- تجهيزات خرج متعددة (مثلاً، مظهار وسماعة أذن).
- منصة يمكنها تنفيذ عدة تطبيقات في الوقت نفسه.
- القدرة على الاتصال لاسلكياً مع أنظمة حقيقة مزيدة أخرى.

نعرض هنا مجموعةً من تحديات الأمن والخصوصية التي ترافق هذه التقانات الجديدة وتطبيقاتها، كما يلخصها الجدول 2. نضع هذه التحديات في محورين: *مدى النظام، والوظيفية*. سندرس في أحد المحورين أنظمة الحقيقة المزيدة المتزايدة المدى: تطبيقات أحادية وتطبيقات متعددة داخل منصة حقيقة مزيدة أحادية، وأنظمة حقيقة مزيدة متعددة التواصل. تظهر التحديات في كل صنف عند مستوى تعقيد النظام أولاً. وعند كل مدى سُنصِّف التحديات فضلاً على ذلك حسب علاقتها بالدخل أو الخرج أو النفاذ إلى المعطيات. نُشجِّع مُصمِّمي تقانات الحقيقة المزيدة المستقبلين على الأخذ بالحسبان تحديات الأمن والخصوصية في هذين المحورين.

يمكن للقراء المُطلعين على أمن الهواتف الذكية أن يلاحظوا بعض التقاطعات بين هذه التحديات والمجموعة التي نعرضها هنا. نشير إلى أنه يمكن تطبيق بعض تقنيات أمن الهواتف الذكية في تقنيات الحقيقة المزيدة؛ في حين تحتاج تقنيات أخرى إلى إعادة نظر ضمن هذا السياق الجديد.

تحديات التطبيقات الأحادية. سندرس أولاً التحديات والتهديدات المحدودة بنطاق تطبيقات الحقيقة المزيدة الأحادية.

الخرج. يجب أن يثق المستخدمون ثقةً كبيرة بتطبيقات الحقيقة المزيدة التي تضع فوق العالم الحقيقي مُدركات مرئية وسمعية ولمسية مع تغذية راجعة افتراضية. يمكن أن تستعمل التطبيقات الخبيثة التجهيزات التي توفر تغذية راجعة غامرة immersive لخداع المستخدمين عما هو في العالم الحقيقي. مثلاً، يمكن لتطبيق خبيث مستقبلي أن يراكب حذاءً خاطئاً للسرعة فوق لافتة حدِّ السرعة الحقيقي (أو أن يضع لافتة وهمية في مكانٍ لا توجد فيه لافتة أصلاً)، أو أن يترجم

نصاً بلغة أجنبية في العالم الحقيقي ترجمةً خاطئة عمداً. عموماً، يمكن أن تخدع مثل هذه التطبيقات المستخدمين ليعتقدوا زيفاً وزوراً أن بعض الأشياء موجودة أو غير موجودة في العالم الحقيقي.

الجدول 2. تحديات الأمن والخصوصية لتقانات الحقيقة المزيدة. صنفنا هذه التحديات في محورين: تحديات مرتبطة بالخرج والدخل والنفوذ إلى المعطيات، كما يحدث في التطبيقات الأحادية، وأنظمة التطبيقات المتعددة والأنظمة المتفاعلة المتعددة.

التطبيقات الأحادية	التطبيقات المتعددة	الأنظمة المتعددة
الخرج	الهجمات المخادعة هجمات الجمل الزائد مسار موثوق للحقيقة	تولّي التضاربات الاتصال بالنقر clickjacking
الدخل	التحقق من الدخل	إضمامة دخل حلّ مسألة التركيز resolving focus
النفوذ إلى المعطيات	التحكم بالنفوذ إلى معطيات المُحس خصوصية المآزة	مشاركة التطبيقات التصالبيهة مشاركة الأنظمة التصالبيهة

يمكن أن تستعمل التطبيقات الخبيثة تقنياتٍ مماثلة لتسبب حِملاً حسيّاً زائداً للمستخدمين. ويمكن للتطبيقات أن تؤذي المستخدمين جسدياً بتوليد ومضاتٍ ضوئية ساطعة على المظهر، أو أصواتٍ عالية أو ردود فعلٍ لمسية شديدة. إن هذه الهجمات ليست غير مسبوقة: فقد استهدف المهاجمون منتديات داء الصرع، فأرسلوا رسائل تحتوي على صورٍ متحركةٍ ومأضة لإثارة أوجاع الرأس أو نوبات الصرع.²⁴ يجب على منصات الحقيقة المزيدة الناشئة أن تأخذ بالحسبان هذه الأنواع من الهجمات وتمنعها.

تكون هذه الهجمات على الخرج أكثر خطورةً في تطبيقات الحقيقة المزيدة الغامرة منها في سيناريوهات الحوسبة المكتبية أو المحمولة باليد الحالية، لأنه من الأصعب على المستخدمين تمييز رد الفعل الافتراضي من الحقيقي، ولأنه من الأصعب على المستخدمين حذف النظام أو توقيفه. كحلٍ أخير لهجمات الخرج، يجب أن يكون المستخدمون قادرين على العودة بسهولة وموثوقية إلى العالم الحقيقي، أي مع التحقق من إيقاف كل أجهزة الخرج.

في المدى القريب، يُشكّل حذف النظام طريقةً بسيطة لتحقيق هذه العودة إلى الحقيقة. مع ذلك، قد يكون صعباً أو مستحيلاً على المستخدمين حذف الأنظمة اللبوسة المستقبلية (مثلاً، العدسات اللاصقة²³ أو التجهيزات المزروعة)، وحتى الأنظمة الحالية اللابوسة يصعب على المستخدمين أن يملصوا منها. مثلاً، أنتج العديد من مُصنّعي السيارات زجاجاً أمامياً يُظهر محتوىً مزيداً فوق ما يراه المستخدم من الطريق.⁵ في هذه الحالات، يجب أن يكون للنظام مساراً موثوقاً يُمكن المستخدم من العودة إلى الحقيقة، شبيه بـ Ctrl-Alt-Del على حواسيب ويندوز. يتطلب تحديد أفضل متتالية مثيلة، أو نمط الدخل الصحيح (مثلاً، حركات أو كلام) إجراء بحثٍ لكل نظام حقيقة مزيدة. توجد مقارنة أخرى تعتمد على حيز منطقة موثوقة من المظهر تُظهر العالم الحقيقي دائماً.

الدخل. ستواجه تطبيقات الحقيقة المزيدة بلا شك تحديات إقرار صلاحية وإبراء sanitization مماثلة لما تواجهه التطبيقات التقليدية. مثلاً، يمكن استغلال تطبيق ترجمة يُحلّل نصاً في العالم الحقيقي بنصٍ وضعٍ بخبث على لافتة. قد يكون تطبيق تقنيات إقرار صلاحية الدخل التقليدية معقولاً في هذه الحالة أيضاً، ولكن يجب أن ينتبه مُصمِّمو أنظمة الحقيقة المزيدة إلى ضرورتها في هذا السياق الجديد.

النفاد إلى المعطيات. تتطلب تطبيقات الحقيقة المزيدة لتوفير وظيفتها المُعدّة لها النفاد إلى مجموعة متنوعة من معطيات المُحسّنات، تشمل البث الفيديوي والصوتي، ومعطيات GPS ودرجة الحرارة وقرارات مقياس التسارع وسواها. سيكون أحد التحديات الهامة لأنظمة الحقيقة المزيدة، كما في أنظمة تشغيل حواسيب سطح المكتب والهواتف الذكية، الموازنة بين النفاد المطلوب للوظيفية وخطر سرقة معطيات التطبيق أو سوء استعمال هذا النفاد. مثلاً، يُمكن لتطبيق خبيث أن يُسرّب موقع المستخدم أو البث الفيديوي لمخدماته الخلفية. يبين هجوم التطبيق PlaceRaider¹⁴، وهو إثبات صحة المفهوم proof-of-concept متوفر في السوق، أنه يمكن استعمال مُحسّنات الهاتف الذكي لتجميع معلومات كافية لتوليد نماذج ثلاثية الأبعاد للبيئات الداخلية.

ستتطلب تطبيقات الحقيقة المزيدة المعقدة، خلافاً لمعظم تطبيقات حواسيب سطح المكتب والهواتف الذكية الحالية، تحسناً دائماً ووافراً. مثلاً، يتطلب التطبيق الذي يكشف ويمسح أوتوماتيكياً الأزمرة (الرمازات) السريعة الاستجابة QR (Quick Response code) النفاد المستمر إلى معطيات الدفقة الفيديوية، مثلما يفعل التطبيق الذي يكشف آلياً إدخال المستخدم كلمة المرور في جهاز آخر ويقدم المساعدة على كتابة كلمة المرور (كما سنناقش لاحقاً). في المحصلة، فإن مخاطر الخصوصية هذه أكبر مما هي في الأنظمة التقليدية.

يجب أن تنتج أنظمة الحقيقة المزيدة مقاربات تُحد من هذه المخاطر. مثلاً، لن تحتاج التطبيقات الفردية على الأغلب النفاد إلى كل معطيات المُحسّن. ربما يتطلب التطبيق النفاد إلى جزء من الشاشة فقط عندما يكون المستخدم في مكان معين، أو يحتاج فقط أن يعرف بعض الأشياء التي يتعرفها النظام (مثلاً، بواسطة أداة تعرف الهياكل Kinect)، ولا يحتاج النفاد إلى كل البث الخام للكَمرة. يجب أن يأخذ مُصمِّمو نظام الحقيقة المزيدة بالحسبان الحُببيّة الملائمة لهذه السماحيات، وسيكون تصميم واجهات تخاطب إدارة السماحية قابلة للاستعمال أمراً هاماً. ومن غير المحتمل نقل scale الحلول الحالية المعتمدة على التقبّل prompt أو القائمة manifest، كما هي مستعملة في الهواتف الذكية، بطريقة قابلة للاستعمال، إذ إنّ احتياجات تطبيقات الحقيقة المزيدة للنفاد الطويل المدى إلى المعطيات (أكثر من مرة واحدة) تجعل تطبيق حلول التحكم بالنفاد في السياق، مثل التحكم بالنفاد المُفوّد بالمستخدم²⁸، تطبيقاً غير مباشر.

تولّد الكَمرات والمُحسّنات الأخرى الشغالة أيضاً خطراً على خصوصية المارة، الذين اعتبرهم كريفلن Krevelen وبولمان Poelman تحدياً للقبول المجتمعي الواسع للحقيقة المزيدة³⁶. يجب أن يكون المارة قادرين على الاختيار بين الظهور في تسجيلات الآخرين أو البقاء مجهولين (مثلاً، غير واضحين)؛ لقد عالج عمل سابق هذه القضايا^{31,39}. قد يحتاج مستعملو الحقيقة المزيدة طرائق يثبتون بها للمارة المُشكّكين أن هذه الضمانات موجودة. يمكن للتشريعات أو لقوى السوق أن تؤدي إلى كَمرات تستجيب لطلبات التجهيزات الأخرى أو المحيط؛ تشير التقارير الإخبارية إلى أن شركة أبل Apple أخذت في الحسبان إضافة هذه القدرة إلى الآيفون iPhone لمنع التصوير الفيديوي للأحداث الحية، مثل الحفلات الموسيقية⁴. يمكن أيضاً أن تنبه الكَمرات المارة أثناء التسجيل، بوميضٍ ضوئي³⁶ أو بتوفير نفادٍ إلى سياسة معلومات أكثر تعقيداً.¹⁹

أتبع مشروع CVDazzle³⁰ مقارنةً مختلفةً- باستعمال ماكياج لتثويش خوارزميات كشف الوجه-توفر الخصوصية دون كمرات متوافقة. تكمن المحدودية الرئيسية في حاجة CVDazzle لضبط يدوي شاق لخوارزمية كشف وجه معين واحد بعينه. أمّا إيجاد خوارزمية عامة لتركيب ماكياج يخدم عملية كشف الوجه فهو مسألة بحثية.

تحديات التطبيقات المتعددة. تُصمّم تطبيقات الحقيقة المزينة وتُنقذ بنماذجها الأولية غالباً على نحوٍ معزول، ومع ذلك نتوقع أن تدعم منصات الحقيقة المزينة المستقبلية، كتلك المبنية على نظارات غوغل أو كينكت مكرسوفت Microsoft Kinect، التطبيقات المتعددة التي تُنفذ في الوقت ذاته، متشاركةً في تجهيزات الدخل والخرج، ومقدمة المعطيات وواجهات تخاطب برامجها Application Program Interface (APIs) بعضها إلى بعض (انظر الشكل 3). يجب أن يستبق الباحثون هذه التطويرات وأن يضمّنوا "نظام تشغيل للحقيقة المزينة" يراعي في تصميمه اعتبارات الأمن والخصوصية الملائمة.



الشكل 3. التطبيقات المتعددة والحقيقة المزينة. ستدعم تطبيقات الحقيقة المزينة البازغة والمستقبلية التطبيقات المتعددة التي تُنفذ في الوقت نفسه، وتتشارك تجهيزات الدخل والخرج، وتعرض المعطيات و APIs بعضها لبعض. في تطبيق متعدد في نظام حقيقة مزينة، ستشارك التطبيقات كتلك المصورة في هذا الشكل الطبيعي تجهيزات الخرج، بما فيها من أجهزة الإظهار والخرج الصوتي والتغذية الراجعة للمسية. يمكن أن يؤدي التضارب بين هذه التطبيقات إلى مخاوف أمنية.

الخرج. في نظام الحقيقة المزينة متعدد التطبيقات، تتشارك التطبيقات تجهيزات الخرج، ومنها أجهزة الإظهار والخرج الصوتي والتغذية الراجعة للمسية. يمكن أن يؤدي التضارب بين التطبيقات المتعددة وهي تحاول استعمال تجهيزات الخرج هذه إلى مخاوف أمنية. مثلاً، يمكن أن يحاول تطبيق خبيث حجب محتوى يقدمه تطبيق آخر (مثال ذلك، تغطية ترجمة صحيحة بصرياً أو سمعياً بترجمة خاطئة).

مع ذلك، ستكون مشاركة الخرج ضروريةً لتوفير الوظيفية المطلوبة لأنظمة الحقيقة المزينة. فمثلاً، قد يرغب المستخدم في رؤية محتوى يغطي رؤيته

للحقيقة في الوقت نفسه من تطبيقات متعددة، مثل الاتجاهات التي يقدمها تطبيق الخرائط، أو بث اجتماعي يُلخص نشاط الأصدقاء القريبين، أو المقطع الذي يعزفه تطبيق موسيقي في نفس اللحظة، إلخ... لذا، فإن الحل الساذج، حيث يتحكم تطبيق واحد فقط بالمظهر في كل مرة (كما في نظام أندرويد Android حالياً، مثلاً) غير كافٍ.

يجب أن تعالج أنظمة الحقيقة المزينة إذن التضارب بين التطبيقات المتعددة عندما تحاول توليد خرج. مثلاً، قد تريد خمسة تطبيقات التعليق على الشيء نفسه (مثلاً، بكتابة الترجمة)، ويحتاج النظام إلى تحديد الأولوية بينها. زيادةً على ذلك، قد تكون معرفة التطبيق الذي أنشأ كل محتوى مهمة للمستخدمين - مثلاً، معرفة فيما إذا كان التعليق الذي يوصي

بمنتج معين قد جاء من صديق أو من مُعلن. يجب أن يُولد مصمّمو نظام الحقيقة المزيدة واجهات تخاطب تجعل مصدر المحتوى الظاهر على الشاشة واضحاً للمستخدمين أو سهل الاكتشاف.

تتطلب الهجمات التقليدية المعتمدة على التلاعب بالخرج مقاربات جديدة أو صياغة جديدة في سياق الحقيقة المزيدة. مثلاً، يمكن للتطبيقات في الأنظمة الحالية أن تُنصّب هجمات clickjacking تخدع المستخدمين بنقّر عناصر حساسة في واجهة التخطاب من تطبيق آخر (مثلاً، ترسل أشياء إلى بروفيل المستخدم في وسائط التواصل الاجتماعي). تعمل هذه الهجمات عموماً إمّا بالتلاعب بمظهر العنصر الحساس - بجعله شفافاً أو بتعتيمه جزئياً بطريقة ذكية - أو بالإظهار المفاجئ للعناصر الحساسة مباشرةً قبل أن ينقر المستخدمون المكان المتوقع. يمكن أن تطوّر التطبيقات المستقبلية لأنظمة الحقيقة المزيدة تقنيات جديدة لخداع المستخدمين بالتفاعل مع العناصر، ويجب على مصممي النظام استباق هذه التهديدات. مثلاً، يمكن لتطبيق الحقيقة المزيدة أن يحاول خداع المستخدم بالتفاعل مع غرض في العالم المادي بدلاً من العالم الافتراضي.

الدخل. من المستبعد أن يتفاعل المستخدمون مع أنظمة الحقيقة المزيدة، باستعمال طرائق الدخل التقليدية مثل نقّر الفأرة أو حتى باستعمال شاشة التحسس باللمس. عوضاً عن ذلك، يمكن أن يزداد تفاعل المستخدمين مع هذه الأنظمة، باستعمال الدخل الدقيق للمُحسّات اللمسية (مثلاً، المُتضمّنة في القفايز)، أو باستعمال الصوت، أو بمساعدة تقانات الملاحقة بالنظر. لن يكون بديهيّاً، مع تقنيات الدخل هذه والتطبيقات المتعددة الجارية، أن يُميّز النظامُ التطبيقَ المركزَ عليه in focus، والذي يجب بالتالي أن يتلقى الدخل.

على سبيل المثال، تحدث التفاعلات الصوتية الحالية إمّا لاحقاً لفعلٍ صريحٍ للمستخدم يدلّ على التطبيق المقصود (مثلاً، نقر الزر "Siri" في الآيفون) أو في الأنظمة التي يمكن لتطبيقٍ واحدٍ فيها فقط أن يستقبل دخلاً صوتياً (مثلاً، في Xbox). عند وجود تطبيقات متعددة فعّالة، ويمكنها استقبال الدخل الصوتي أو غيره في لحظة معينة، فإنه يجب أن توجد طريقةً صالحةً للاستعمال للمستخدمين لجلب التطبيقات إلى بؤرة التركيز، أو للنظام لتحديد الوجهة الصحيحة المقصودة لأوامر الدخل عندما تكون البؤرة مُلتبسة. نوّكد أن أنظمة الحقيقة المزيدة المستقبلية ستتقدّم على الأرجح تطبيقات متعددة في الوقت نفسه، وسيُنفذ العديد منها وينصت للدخل، دون أن يكون لديها أي خرج مرئي. قد يُسهّل التصميم غير الصحيح لتمييز البؤرة عمل التطبيقات الخبيثة في سرقة دخل المستخدم الموجّه لتطبيقٍ آخر (مثلاً، لسرقة كلمة مرور موجّهة إلى صندوق طلب الدخول لتطبيقٍ آخر). مثلاً، يمكن أن يحاول تطبيق خبيث تسجيل كلمة مفتاحية لفظية لها رنة صوت مشابهة، لتطبيقٍ آخر حساس، مما يزيد عمداً غموض الدخل.

النفاز إلى المعطيات. سترغب تطبيقات الحقيقة المزيدة، على الأرجح، كما في أنظمة التشغيل التقليدية، في عرض واجهاتٍ تخاطب البرامج بعضها لبعض، وقد يرغب المستخدمون في مشاركة الأغراض الافتراضية بين التطبيقات. يجب أن يستكشف الباحثون نماذج تحكم بالنفاز إلى المعطيات ملائمة للمشاركة بين التطبيقات المتصالية. يمكن بالتأكيد تطبيق الدروس المأخوذة من التصميم التقليدي للتحكم بالنفاز في هذا الفضاء، ولكن قد تتطلب التقانات والبيئات الحديثة مقاربات جديدة. مثلاً، النسخ واللصق والسحب والإفلات حركاتٌ للمستخدم متعارفة لمشاركة المعطيات بين التطبيقات التقليدية، ومن ثم لها تبعات على التحكم بالنفاز. حاولت سلسلة طويلة من الأعمال في أنظمة حواسيب سطح المكتب والهواتف الذكية أن تقابل أفعال المستخدم بامتيازاته في التطبيق (تتضمن الأمثلة Miller²¹ و Roesner et al.²⁸)؛ ستحتاج أنظمة الحقيقة المزيدة إلى تطوير حركات مستخدم جديدة للدلالة على الرغبة في التشارك. إضافةً إلى ذلك، من المستبعد

أن تُظهر أنظمة الحقيقة المزيدة التطبيقات في نوافذ مُبوبة labeled بالطريقة التي تُظهرها أنظمة تشغيل حواسيب سطح المكتب التقليدية، لذا فهناك حاجة لمقاييس تفاعل جديدة لتمكين المستخدمين من تحديد التطبيقات والدلالة على التطبيق الذي يجب أن يتلقى المعطيات المشتركة.

تحديات الأنظمة المتعددة. بعد عرض نظام الحقيقة المزيدة الأحادي الذي يُنفذ تطبيقات متعددة، سنعرض التفاعلات بين أنظمة الحقيقة المزيدة المتعددة العائدة لمستخدمين مختلفين. تعرض الأعمال السابقة في مجال الحقيقة المزيدة تطبيقات تعاونية بين مستخدمين متعددين لنظام الحقيقة المزيدة. تتضمن هذه التطبيقات ألعاباً متعددة اللاعبين^{11 و20 و40}، وحضور المؤتمرات من بُعد¹⁶، والتعاون وجهاً لوجه²⁶. يطرح هذا النوع من التطبيقات تحديات إضافية في الأمن والخصوصية.

الخرج. قد يكون للمستخدمين المتعددين رؤى مختلفة للعالم الذي تعرضه أنظمة الحقيقة المزيدة الخاصة بهم. مثلاً، يمكن أن يرى المستخدمون المختلفون إعلانات افتراضية مختلفة متراكبة على لوحات العالم الحقيقي، أو يمكن أن تُعرض محتويات مختلفة على مستخدمين مختلفين يشاهدون عرضاً ما حسب مستويات نفاذهم (مثلاً، يمكن أن يرى أحد المستخدمين الحواشي البالغة السرية، في حين لا يستطيع الآخرون ذلك). تتطلب هذه الرؤى المتضاربة أن يدير المستخدمون نماذج ذهنية لمن يمكنه ملاحظة هذه المعلومات أو تلك، خشية كشف معلومات شخصية تخصهم وحدهم، من غير قصد. تتطلب معالجة هذه القضية إبداعات في تصميم واجهة التخاطب لمساعدة المستخدمين بهذه المهمة.

الدخل. يرتبط ازدياد تعقيد أنظمة الحقيقة المزيدة وتطبيقاتها ارتباطاً وثيقاً بازدياد عدد دخل المُحسّات وتعقيدها التي توفرها التقانات التمكينية. بدورها، ستؤدي هذه الوفرة في دخل المُحس من العديد من المستخدمين إلى تطبيقات تحسّس تعاونية، يمكنها ذاتها إعادة تغذية تطبيقات الحقيقة المزيدة بالمعطيات. مثلاً، تستعمل غوغل بالفعل المعطيات التي تجمعها هواتف المستخدمين الذكية لتقدير ظروف حركة المرور، لتعيدها بعد ذلك إلى هواتف المستخدمين⁸. هذا النوع من المعطيات ضروري لتمكين تطبيقات الحقيقة المزيدة المستقبلية من الإظهار على الزجاج الأمامي للسيارة، مثلاً.

بيد أنه يمكن أن يستعمل المستخدمون الخبثاء هذا النوع من إضامات الدخل لخداع أنظمة جميع المعطيات. فمثلاً، يمكن أن يستفيد موقع مراجعة review site من تعقب الموضع لقياس شعبية مطعم بملاحظة متوسط عدد الأشخاص الموجودين فيه أثناء اليوم. يمكن عندئذٍ لصاحب مطعم حكيم أن يدفع للناس أجراً ليقفوا في المطعم دون أن يشتروا شيئاً. تزداد شعبية المطعم المقيسة، دون أن يكون لها أي علاقة بجودته.

ستقود تقانات الحقيقة المزيدة التي تُجمع المعطيات باستمرار إلى اعتماد تطبيقات التحسّس التعاونية هذه؛ لذا، ستزداد أهمية مخاوف الأمن. مثالٌ آخر: تُجمع شبكة جماعة الزلازل Community Seismic Network معطيات مُحسّات قياس التسارع لعدة أشخاص للكشف عن الهزات الأرضية والتنبؤ بها؛ يمكن أن يتلاعب مهاجمٌ بالمحسّات، لإجراء محاكاة مزيفة لنشاط زلزالي غير اعتيادي، مثلاً بتشجيع عدة أشخاص مسجلين في المشروع على القفز دفعةً واحدة في سياق لعبة لا علاقة لها بالأمر. (مثلاً، تطلب لعبة Improv Everything¹³ من المستخدمين تشغيل الملفات الصوتية التي توفرها لهم في الوقت المُحدّد واتباع التعليمات الصوتية). لا تساعد المُحسّات الموثوقة³⁹ في هذه الحالات - مع أنها هامة لتقادي الهجمات الأخرى - لأنه جرى التلاعب بظروف العالم الحقيقي.

النفاذ إلى المعطيات. ستمكّن أنظمة تواصل الحقيقة المزيدة المستخدمين من تشارك المحتوى الافتراضي، إضافةً إلى إظهار مختلف المحتويات لمختلف المستخدمين. مثلاً، يمكن أن يُنشئ مستخدمٌ وثيقةً افتراضيةً في نظام

الحقيقة المزيدة الخاص به، وأن يُقرّر لاحقاً مشاركة إظهارها في أنظمة المستخدمين الآخرين. قد تكون بعض المشاركات ضمنية؛ لنتخيل نظام حقيقة مزيدة يستعمل آلياً تسجيلات كميرات المستخدمين القريبين منه لتزويد مستخدم معين بنموذج ثلاثي الأبعاد عنه أو عنها، في الزمن الحقيقي.

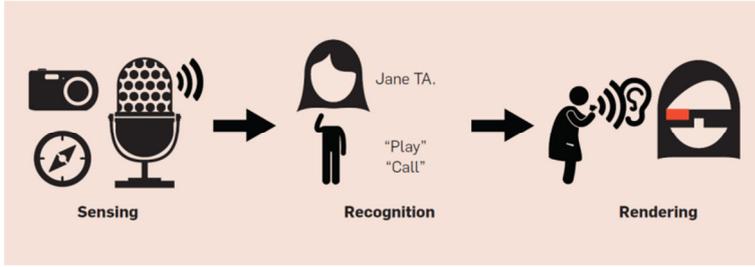
يمكن للمشاركة الضمنية أو الصريحة للمعطيات بين أنظمة الحقيقة المزيدة المنفصلة أن تُمكن تطبيقات قيّمة. مع ذلك، نحتاج إلى نماذج وواجهات تخاطب ملائمة للتحكم بالنفاذ لتمكين إدارة هذه المشاركة. حالياً، يعاني المستخدمون فعلياً من صعوبات في تشكيل نماذج ذهنية لإعدادات خصوصياتهم في خدمات مثل فيسبوك Facebook بسبب تعقيد العلاقات بين الأشخاص وبنود المعطيات.¹⁸ ستجعل كمية المعطيات الهائلة التي تجمعها أنظمة الحقيقة المزيدة ومكاملة الأشياء الافتراضية مع العالم الحقيقي هذه المسألة أكثر صعوبة.

التوجّهات الدفاعية

سنعرض هنا الخطوط العريضة للتوجّهات الدفاعية لتقانات الحقيقة المزيدة. بدايةً، تشبه بعض تحديات الأمن والخصوصية المرافقة لتقانات الحقيقة المزيدة تلك التي تواجهها حالياً الهواتف الذكية، مثل خصوصية معطيات المُحس ومشاركة التطبيقات التصلابية. في بعض الحالات، يتمثل أحد التوجّهات الدفاعية الملائمة للحقيقة المزيدة في اعتماد حلول الهواتف الذكية. مثلاً، يمكن في المدى القصير اعتماد كشف السماح وإجرائية مراجعة مخزن التطبيقات.

مع ذلك، توجد في المدى البعيد، أسبابٌ متعددة توجب اختلاف المقاربات في سياق الحقيقة المزيدة عن حلول الهواتف الذكية. أولاً، يُبين تحليل حاجات تطبيقات الهواتف الذكية من الموارد²⁸ أن معظم التطبيقات تتطلب النفاذ إلى معظم الموارد مرةً واحدةً أو لمدى قصير، مما يجعل الحلول التي تتطلب تفاعلات المستخدم في السياق (مثل التحكم بالنفاذ المُفود بالمستخدم²⁸) ممكنةً. على النقيض من ذلك، ستطلب تطبيقات الحقيقة المزيدة نفاذاً دائماً أو بعيد المدى إلى معطيات المُحس يتجاوز نطاق تطبيقات الهواتف الذكية. إضافةً إلى ذلك، لن تكون موارد الحقيقة المزيدة واضحةً للمستخدمين وللمارةً وضوح الموارد في سياق الهاتف الذكي - مثلاً، ستكون كمرّة نظام الحقيقة المزيدة دائماً شغالة، في حين توفر كمرّة الهاتف الذكي، حتى عندما تُشغّلها البرمجيات الخبيثة، معطيات أقل بكثير أثناء وجود الهاتف في جيب المستخدم. لذا، فإننا نرى من المهم الأخذ في الحسبان سياقات الحقيقة المزيدة المستقبلية بكل معنى الكلمة عند تصميم الحلول في هذا الفضاء.

على هذا المنوال، سيحتاج الأمر إلى بحوثٍ جديدة في الحلول الخاصة بالحقيقة المزيدة. مثلاً، بدأ الباحثون بالأخذ في الحسبان دعمَ نظام التشغيل الخاص بالحقيقة المزيدة.⁷ تتبع تطبيقات الحقيقة المزيدة - ونظام التشغيل التحتي - طبيعياً التوارد pipeline المبين في الشكل 4، ومن ثمّ يمكن توصيف البحوث وفقاً لذلك، ويمكن أن تفترض النماذج المختلفة للبحث حدوداً مختلفة بين التطبيق ونظام التشغيل. في المرحلة الأولى، التحسُّس، يجمع التطبيق (أو نظام التشغيل) المعطيات الحسيّة الأولية مثل الصوت أو الفيديو أو الأمواج الراديوية؛ تشمل البحوث هنا حدّاً لما يتحسُّس المعلومات التي تُجمَع (مثلاً، الكميرات اللطيفة³¹) أو حدّاً لاستعماله (مثلاً، سياسات الاستبقاء). ثانياً، في مرحلة التعرّف، تستخلص خوارزميات تعلم الآلة الأغراض بدلالة semantics عالية المستوى: مثلاً على ذلك، يبين الشكل هيكل Kinect، ووجهاً، والاسم المرافق، ومُثيرات الأمر الصوتي. تشمل البحوث ذات الصلة أغراضاً متغيرة تسبب سلبيات زائفة (مثلاً، CVDazzle¹⁰)، والسياسات التي تحكم نفاذ التطبيقات إلى الأغراض¹⁵. أخيراً، تصيير التطبيق (أو نظام التشغيل) فوق



الشكل 4. قناة توارد الحقيقة المزيدة. تقوم تطبيقات الحقيقة المزيدة (1) بتجميع معطيات المُحسّات، (2) وتستخرج منها الأغراض بمستوى دلالي عالٍ. أخيراً، تضعها (3) فوق حواس المستخدم.

حواس المستخدم، مثل الرؤية والسمع. تشمل البحوث هنا كشف اللامتغيرات التي يجب احترامها لتفادي إيذاء المستخدم ولبناء "تصييرٍ موثوق" جيد الأداء يحترم هذه اللامتغيرات.

لن تقوم كل التوجهات الدفاعية للحقيقة المزيدة على الحلول التقنية. قد تستدعي بعض التحديات مقاربات اجتماعية أو سياسية أو قانونية؛ مثلاً، السياسات الممكنة لتعطيل الماّرة والكميرات

المتوافقة، كما ناقشناه سابقاً. ستستفيد القضايا الأخرى بالمثل من المقاربات اللاتقنية.

أخيراً، نطالب الباحثين العاملين في هذا المجال بمرحلة تجريبية. تعتمد معظم تطبيقات الحقيقة المزيدة التجريبية حالياً على كينكت مكرسوفت أو منصات الهواتف الذكية مثل Layar؛ وكلاهما يستدعي تطبيقاتٍ فردية فقط تُنفَّذ في آنٍ واحد، ومن ثم تُخفي التحديات التي تظهر عند ازدياد تعقيد أنظمة الحقيقة المزيدة.

تطبيقات جديدة

مع أن تقانات الحقيقة المزيدة تخلق مخاوف كبيرة في الأمن والخصوصية، فإن لها فرصة غير مُستكشفة لتعزيز الأمن والخصوصية، بتطبيقها على مسائل قائمة. سنناقش هنا فرص التطبيقات المعززة للأمن والخصوصية التي تُمكنها تقانات الحقيقة المزيدة وأنظمتها. لا ريب أن قائمتنا غير كاملة؛ ونأمل أن نرى عملاً مستقبلياً غنياً في هذا المجال.

تعزيز وجهات النظر الشخصية. يمكن أن تُعزّز أنظمة الحقيقة المزيدة التي تدمج أجهزة الإظهار الشخصية الرأسية heads-up أو غيرها من هذه الرؤى الشخصية لمعالجة قضايا الأمن والخصوصية الراهنة- على وجه الخصوص، حماية المعطيات الخاصة وتحسين إدارة كلمة المرور.

تتمتع أجهزة الإظهار الشخصية بدفاعٍ قوي مُضاد للتجسس بالتلصص من فوق الأكتاف shoulder surfing، لأنه يمكن للمستخدمين أن يتفاعلوا مع التطبيقات المرئية في مجال رؤيتهم فقط. على سبيل المثال، يكشف مستخدم الحاسوب المحمول في طائرة حالياً كل شيء يراه أو يكتبه للجالسين بجواره، وقد برهن الباحثون أنه يمكن استعمال لقطات من كمرّة رخيصة الثمن لإعادة بناء ما يكتبه المستخدم على لوحة مفاتيح افتراضية نقّالة.²⁵ يُمكن مظهر شخصي رأسي مدمج مع مُحسّ لمسي للإدخال السري من تحسين الخصوصية تحسناً كبيراً.²

² نلاحظ أن أجهزة الإظهار التي يمكن الرؤية من خلالها، كذلك التي تستعملها نظارات غوغل، لا يمكن أن تتمتع بالخصوصية الشاملة من المراقبين الخارجيين. على سبيل المثال، يمكن استعمال الصور المأخوذة من أجهزة الإظهار باستعمال عدسة مقربة لإعادة بناء محتوى الشاشة، على نحوٍ مماثل لإعادة البناء من انعكاسات الشاشة.² يجب أن توصّف البحوث المستقبلية هذا التهديد توصيفاً شاملاً وأن تُصمّم الدفاعات الملائمة.

زيادةً على ذلك، تُمكن أجهزة الإظهار الشخصية المحتوى المُعمى في العالم الحقيقي الذي يمكن فك تعميته بأنظمة الحقيقة المزيدة للمستقبلين المقصودين فقط. على سبيل المثال، يمكن لشركة أن ترسل إشعارات مُعماة إلى لوحة نشرات bulletin board يمكن أن يقرأها العاملون بواسطة أنظمة الحقيقة المزيدة الصادرة عن شركتهم، ولكن لا يمكن لزائري بناء الشركة قراءتها، (يتطلب تخزين المفتاح فقط، وليس المحتوى المُعمى، في مخدّم يمكن النفاذ إليه بنظام الحقيقة المزيدة من الخصوم العثور على الإشعار الفيزيائي، بدلاً من مجرد افتضاح مخدّمات الشركة.) إن بشائر مثل هذه الأنظمة ممكنة حالياً باستعمال الهواتف الذكية والأرمزة القضبانبة الثنائية البعد التي تُرمز عناوين الموارد النظامية URLs بمعطيات ذات تحكم بالنفاذ ملائم؛ ستُجنّب أجهزة الإظهار الرأسية المزيدة الحاجةً للمسح اليدوي.



الشكل 5. نموذج أولي لمدير كلمات مرور يستعمل الحقيقة المزيدة. يُظهر توسع Chrome (الخلفية) الذي أنجزناه رمز QR يُمثل موقع الويب الحالي. يسمح تطبيقنا لنظارات غوغل (الوجهية)، استجابةً للأمر الصوتي "find password" هذا الرمز QR ويُظهر كلمة المرور المُخزّنة لذلك الموقع على نحوٍ سري في مظهر الرأس.

يمكن لأنظمة الحقيقة المزيدة أن تقوم أيضاً بدور مدير مُعزّز لكلمات مرور المستخدمين، بتقديم كلمات مرور أو تلميحات لكلمة المرور بواسطة الإظهار الشخصي. على سبيل المثال، يمكن لمظهر أن يرسم outline المحارف الملائمة التي يجب أن يُدخلها المستخدم في التجهيزات القديمة مثل لوحات ATM PIN. يمكن إعطاء المستخدمين عندئذٍ كلمات مرور معقدة، لأنهم لن يحتاجوا فعلياً لحفظها أبداً. يتطلب هذا التطبيق تعقّباً بلا أثر وتصميم نظام يحمي كلمات المرور المُخزّنة بشكلٍ صحيح.

مثال ملموس: أنجزنا نموذجاً أولياً لتطبيق يدير كلمات المرور، يتكون من تطبيق نظارات غوغل وتوسع متصفح (Chrome) (انظر الشكل 5). يغير توسع (Chrome) واجهة تخاطب

المستخدم للمتصفح لتظهر رمز QR يُمثل موقع الويب الذي يظهر للمستخدم في حينه (موقع الويب في شريط عنوان المتصفح). يمكن أن يطلب المستخدمون من تطبيق نظارات غوغل مسح هذه الأرمزة QR وتستشير قاعدة معطيات كلمة مروره باستعمال الأمر الصوتي "Ok Glass, find password". إذا كان المستخدم قد خزّن سابقاً كلمة مرور لموقع الويب ذلك، فإن التطبيق يُظهر كلمة المرور؛ بخلاف ذلك، يمكن أن يُسجّل المستخدم كلمة مرور جديدة بسؤال توسع Chrome توليد رمز تسجيل QR ويطلب من Glass تخزين كلمة المرور الجديدة باستعمال الأمر الصوتي "enroll password". جعلنا رمز نموذجنا الأولي متاحاً في الموقع: <https://github.com/froeschele/GlassPass>.

بتصميم الرماز QR الذي يُظهره توسع المتصفح بحيث يتضمن سراً مشتركاً بين المتصفح والهاتف، يمكن لهذا التطبيق أن يقوم ، زيادةً على ذلك، بالحماية من التزوير، لأن مواقع الوب غير قادرة على توليد وإظهار أرمزة QR المزورة التي تطابق كلمات مرور شرعية في مدير كلمة المرور.

تعزيز أنظمة المُحسّات المعقّدة. تُعزّز أنظمة الحقيقة المزيدة ضمّ العديد من المداخل وتجهيزات التحسس، التي يمكن ضمّها لتعزيز الأمن المادي والرقمي والخصوصية والأمان.

يمكن أن تُعزّز الأنظمة المستقبلية تقانات الحقيقة المزيدة للكشف عن شروط الخصوصية والأمن التي يجب تنبيه المستخدم عنها. على سبيل المثال، بدلاً من الاعتماد على كمرات متوافقة compliant لحجب المستخدمين من التسجيل غير المرغوب فيه، يمكن للنظام تنبيه المستخدمين عندما يكشف عدسات كمرّة توجّه إليهم، مستعملاً (على سبيل المثال) الرؤية الحاسوبية للكشف عن بريق الضوء المنعكس عن العدسة³⁵. كما يمكنه أن يكشف أيضاً عن بعض أشكال التتصت، على سبيل المثال، مكرووفون ليزري موجّه نحو نافذة.

يمكن أن تكشف مثل هذه الأنظمة أيضاً عن محاولات غش مادية. على سبيل المثال، يمكن أن يُقدّر نظام الحقيقة المزيدة أبعاد وشكل شقّب slot لوحة ATM، وعندما يُصدر تحذيراً إذا تبين أنه جرى إضافة تجهيزة تزوير لوحة. بالمثل، يمكن تطبيق الأعمال الحالية عن التفسير الحاسوبي لتعبيرات الوجه¹² للكشف عن الكذب المعتمد على السلوك.³⁸ يشير أحد زملائنا إلى هذا التطبيق على أنه "حس التجسس".

يمكن استعمال أنظمة الحقيقة المزيدة، زيادةً على تخزين كلمات المرور، للاستيقان الضمني لمستخدميها. يمكن استعمال العدد الكبير من المُحسّات المرفقة بالأشخاص الذين يستعملون هذه التقانات لاستيثاقهم عن طريق الخصائص الحيوية والسلوكية. اختبر عمل سابق إمكان مثل هذه الآليات في الهواتف النقّالة^{14,27}. يمكن لأنظمة الحقيقة المزيدة أن توفر استيثاقاً أقوى بكثير. بالمثل، يمكن استعمال معطيات المُحسّ للمساعدة في قرارات التحكم بالتحويل والنفاد. زيادةً على المُحسّات المتعلقة بشخص معين (مثلاً، السيدة أليس)، يمكن أيضاً استعمال مُحسّات المارّة لاستيثاقها بتزويد نظام الاستيثاق بمشاهد بصرية وسمعية وتحسسية أخرى للسيدة أليس، من طرف ثالث. يوزع نظام الاستيثاق من طرف ثالث هذا الثقة للأنظمة والأشخاص دون أن يكون لديه أي دافع لاستيثاق مزورّ للسيدة أليس.

الخاتمة

تمتلك أنظمة الحقيقة المزيدة، مع مداخلها ومخارجها المتقدمة والمنتشرة وقدراتها على المعالجة، القدرة على نفع عدد كبير من المستخدمين منفعةً كبيرة. لاستكمال الإبداعات الجارية في تقانات الحقيقة المزيدة، نعتقد أنه قد آن الأوان لتعريف خريطة طريق لحماية أمن وخصوصية الحاسوب لأنظمة الحقيقة المزيدة- قبل أن تصبح هذه الأنظمة واسعة الانتشار وتصبح بنياناتها راسخة. لتحفيز خريطة الطريق هذه، ندرس التحديات الجديدة للأمن والخصوصية التي تطرحها هذه الأنظمة، ونستكشف الفرص التي تتيحها هذه التقانات لتوليد تطبيقات جديدة تُعزّز الخصوصية والأمن.

المراجع

- [1] Azuma, R.T. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments* 6 (1997), 355–385.

- [2] Azuma, R., Bailiot, Y., Behringer, R., Feiner, S., Julier, S. and Macintyre, B. Recent advances in augmented reality. *IEEE Computer Graphics and Applications* 21, 6 (2001), 34–47.
- [3] Backes, M., Chen, T., Duermuth, M., Lensch, H. and Welk, M. Tempest in a teapot: Compromising reflections revisited. *IEEE Symposium on Security and Privacy* (2009).
- [4] Business Insider. This apple patent will shut down your camera at live concerts; <http://www.businessinsider.com/iphone-concert-patent-2011-6>.
- [5] CNN. Augmented-reality windshields and the future of driving, 2012; <http://virtual.vtt.fi/virtual/proj2/multimedia/alvar.html>.
- [6] Costanza, E., Kunz, A. and Fjeld, M. Mixed reality: A survey. *Human Machine Interaction*. Springer-Verlag, 2009, 47–68.
- [7] D’Antoni, L., Dunn, A., Jana, S., et al. Operating system support for augmented reality applications. In *Proceedings of USENIX Workshop on Hot Topics in Operating Systems* (2013).
- [8] Google. Crowdsourcing road congestion data; <http://googleblog.blogspot.com/2009/08/bright-side-of-sitting-in-traffic.html>.
- [9] Halderman, J.A., Waters, B. and Felten, E.W. Privacy management for portable recording devices. In *Proceedings of the 3rd ACM Workshop on Privacy in Electronic Society* (2004).
- [10] Harvey, A. CVDazzle: Camouflage from Computer Vision; <http://cvdazzle.com/>.
- [11] Henrysson, A., Billinghurst, M., And Ollila, M. Face to face collaborative AR on mobile phones. In *Proceeding of the 4th IEEE/ACM International Symposium on Mixed & Augmented Reality* (2005).
- [12] Hoque, M.E., McDuff, D. and Picard, R.W. Exploring temporal patterns in classifying frustrated and delighted smiles. *IEEE Transactions on Affective Computing* 3 (2012), 323–334.
- [13] Improv Everywhere. The Mp3 Experiments, 2012; <http://improveverywhere.com/missions/the-mp3-experiments/>.
- [14] Jakobsson, M., Shi, E., Golle, P., and Chow, R. Implicit authentication for mobile devices. In *Proceedings of the 4th USENIX Workshop on Hot Topics in Security* (2009), USENIX.
- [15] Jana, S., Molnar, D., Moshchuk, A. et al. Enabling fine-grained permissions for augmented reality applications with recognizers. Tech. Rep. MSRTR-2013-11, Microsoft Research, Feb. 2013.
- [16] Kato, H. and Billinghurst, M. Marker tracking and HMD calibration for a video-based augmented reality conferencing system. In *IEEE/ACM Workshop on Augmented Reality* (1999).
- [17] Laycock, S. and Day, A. A survey of haptic rendering techniques. *Comp. Graphics Forum*. 26, 1 (2007), 50–65.
- [18] Madejski, M., Johnson, M. and Bellovin, S.M. The Failure of Online Social Network Privacy Settings. Tech. Rep. CUCS-010-11, Dept. of Comp. Science, Columbia University, 2011.
- [19] Maganis, G., Jung, J., Kohno, T. et al. Sensor Tricorder: What does that sensor know about me? In *Proceedings of the 12th Workshop on Mobile Computing Systems and Applications* (2011), ACM.
- [20] Microsoft. Kinect for Windows, 2012; <http://www.microsoft.com/en-us/kinectforwindows/>.
- [21] Miller, M.S. Robust Composition: Towards a Unified Approach to Access Control and Concurrency Control. Ph.D. thesis. Johns Hopkins University, Baltimore, MD, 2006.
- [22] Papagiannakis, G., Singh, G. and Magnenatthalmann, N. A survey of mobile and wireless technologies for augmented reality systems. *Computer Animation and Virtual Worlds* 19 (2008), 3–22.
- [23] Parviz, B. For your eye only. *IEEE Spectrum* 46 (2009), 36–41.
- [24] Poulsen, K. Hackers assault epilepsy patients via computer. *Wired* (2008); <http://www.wired.com/politics/security/news/2008/03/epilepsy>.

- [25] Raguram, R., White, A.M., Goswami, D. et al. iSpy: automatic reconstruction of typed input from compromising reflections. In *Proceedings of the 18th ACM Conf. Computer and Communications Security*.
- [26] Reitmayr, G. and Schmalstieg, D. Mobile collaborative augmented reality. In *Proceedings of the 4th International Symp. on Augmented Reality* (2001).
- [27] Riva, O., Qin, C., Strauss, K., And Lymberopoulos, D. Progressive authentication: Deciding when to authenticate on mobile phones. In *Proceedings of the 21st USENIX Security Symposium* (2012).
- [28] Roesner, F., Kohno, T., Moshchuk, A. et al. User-driven access control: Rethinking permission granting in modern operating systems. *IEEE Symposium on Security and Privacy* (2012).
- [29] Saponas, T.S., Tan, D.S., Morris, D. et al. Enabling always-available input with muscle-computer interfaces. In *Proceedings of the 22nd ACM Symposium on User Interface Software and Technology* (2009).
- [30] Saroiu, S. and Wolman, A. I am a sensor, and I approve this message. In *Proceedings of the 11th Workshop on Mobile Computing Systems and Applications* (2010), ACM.
- [31] Schiff, J., Meingast, M., Mulligan, D.K., Sastry, S. and Goldberg, K.Y. Respectful cameras: Detecting visual markers in real-time to address privacy concerns. In *Proceeding of the 2007 Int'l Conference on Intelligent Robots and Systems*.
- [32] Starner, T., Leibe, B., Singletary, B. and Pair, J. Mindwarping: Towards creating a compelling collaborative augmented reality game. In *ACM Intelligent User Interfaces* (2000).
- [33] Sutherland, I.E. A head-mounted three-dimensional display. In *Proceedings of the Fall Joint Computer Conference, American Federation of Information Processing Societies* (1968).
- [34] Templeman, R., Rahman, Z., Crandall, D.J., And Kapadia, A. Placeraider: Virtual theft in physical spaces with smartphones. CoRR abs/1209.5982 (2012).
- [35] Truong, K., Patel, S., Summet, J. and Abowd, G. Preventing camera recording by designing a capture-resistant environment. *Proceedings of Ubicomp* (2005).
- [36] Van Krevelen, D. and Poelman, R. A survey of augmented reality technologies, applications, and limitations. *The International Journal of Virtual Reality* 9 (2010), 1–20.
- [37] Viola, P. and Jones, M. Robust real-time object detection. *International Journal of Computer Vision* 57 (2004), 137–154, Hingham, MA.
- [38] Vrij, A., Edward, K., Roberts, K. and Bull, R. Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal Behavior* 24 (2000), 239–263.
- [39] VTT Technical Research Centre Of Finland. Alvar Software Library, 2009. <http://cnn.com/2012/01/13/tech/innovation/ces-future-driving/>.
- [40] Wagner, D., Pintaric, T., Ledermann, F. and Schmalstieg, D. Towards massively multi-user augmented reality on handheld devices. In *Proceedings of the 3rd International Conference on Pervasive Computing* (2005).
- [41] Zhou, F., Duh, H. B.-L. and Billinghurst, M. Trends in augmented reality tracking, interaction and display: A review of 10 years of ISMAR. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality* (2008).

المخاطر والأساطير المتعلقة بالحوسبة

السحابية والتخزين السحابي

RISKS AND MYTHS OF CLOUD COMPUTING AND CLOUD STORAGE*

Peter G. Neumann

ترجمة: م. سماح راغب
مراجعة: د. يُمن أناسي

دراسة الأنماط الحالية والجديدة للمخاطر الكامنة في الخدمات السحابية.

يُنظر إلى الحوسبة والتخزين السحابيين غالباً على أنهما نعمتان عامتان، إن لم تكونا خلاصاً مالياً. وهناك أسباب وجيهة تكمن وراء هذا الادعاء. فالخدمات السحابية تكون عادة على أرض الواقع أرخص بكثير من نظيراتها المخصصة لذلك، حيث تكون الإدارة والرقابة الإدارية أبسط في ظل سلطة مركزية واحدة. استفادت الشركات الصغيرة والشركات الناشئة من استعمال الخدمات السحابية المنخفضة التكلفة خلال سنواتها القليلة الأولى. تُعد المنصات السحابية سُبلاً مهمة للعديد من الشركات للبدء بنشاطها، فهي تتيح لها الوصول إلى العديد من الزبائن بتكلفة منخفضة، وينظر الكثير من رجال الأعمال إلى السحابية على أنها محرك للشركات الصغيرة ولخلق فرص العمل.

تُعتبر خدمات التخزين السحابية أيضاً هبةً للمستخدمين الأفراد الذين لا يحتفظ معظمهم بانتظام بنسخة احتياطية لأجهزة الحاسوب والأجهزة النقلة الخاصة بهم، أو لا يقومون بذلك على الإطلاق. يعمل النسخ الاحتياطي المنخفض التكلفة والتلقائي ضمن سحابة التخزين على حماية معطياتهم القيمة من الضياع.

مع كل هذه المزايا، تُبدي الخدمات السحابية أيضاً أنواعاً جديدة من المخاطر التي سنعالجها هنا. ينبغي على المستخدمين المحتملين للسحابة تقييم هذه المخاطر قبل اتخاذ قراراتهم بشأن كيفية استعمال السحابات. تتمثل القضية الرئيسية في أن توقعات أهلية السحابات للثقة قد تكون غير واقعية، إذ قد تكون السرية، وسلامة النظام، وسلامة المعطيات، والموثوقية، والمتانة، والمرونة موضع شك. تُعتبر الحماية من كلٍّ من المراقبة والحرمان من الخدمة، إضافة إلى الوصول المستمر وتوافقية المعطيات المخزنة على المدى الطويل قضايا ذات أهمية. كما يجب النظر أيضاً في نزاهة، وقابلية محاسبة، وصدقية أطراف ثالثة بل حتى أطراف عديدة أخرى غير معروفة يُحتمل أن تكون غير جديرة بالثقة. قد

* تُشير هذا البحث في مجلة *Communications of the ACM*، المجلد 57، العدد 10، تشرين الأول (أكتوبر) 2014، الصفحات 25 – 27.

تمتلك تلك الأطراف نماذج أعمال تتعارض جذرياً مع احتياجات المستخدمين، ومن ناحية أخرى، قد تتعرض للإفلاس ويُلقي باللائمة على المستخدمين. فضلاً على ذلك، فإن إساءة الاستعمال الداخلية قد تؤدي إلى مخاطر إضافية. تتعلق كل هذه المخاطر بأنواع كثيرة مختلفة من التطبيقات، وكمثال على ذلك، قد يكون البريد الإلكتروني غير المعمى الذي يحتفظ به مقدّم السحابة على وجه الخصوص محفوفاً بالمخاطر بحسب وجهات نظر المستخدمين.

يعود ظهور المفهوم الأساسي للحوسبة والتخزين السحابيين إلى جيلين سابقين يزخران بخبرة كبيرة في تصميم وإدارة هذه النظم. اكتسب منذ ستينيات القرن الماضي الكثير من الخبرات المتعلقة بنظم اقتسام الزمن وموارد الحوسبة المشتركة وإمكانات النفاذ التعاوني للمعطيات (Tymshare، Multics، CTSS)، وشملت أنواعاً مختلفة من المشاركة. وقد وظّف مشروع Athena في معهد ماساتشوستس للتقانة MIT نظام ملفات شبكة Sun منذ 1980s، واستعمل بعد ذلك نظام الملفات Andrew لتقديم ثلاث خدمات يمكن تحديدها بحسب الخدمات السحابية المتاحة اليوم بـ: تخزين برامج التطبيقات عن بُعد، وتخزين الملفات الشخصية عن بُعد، والنسخ الاحتياطي عن بُعد للملفات الشخصية. ومع ذلك، خضع اقتسام الزمن وخدمات Athena الثلاثة لإدارةٍ عملية وحيدة، وهذا ما قلل من عدد الكيانات التي يجب على المستخدمين الوثوق بها (وأتاح في الوقت نفسه نقطة فشل وحيدة). نعلم من خلال التجربة كيف يجري تعويض بعض المخاطر عندما تُلقى مسؤولية الخدمات السحابية على عاتق إدارة واحدة. يمكن التقليل من بعض المخاطر المتعلقة بالخدمات المحلية والبعيدة، فعلى سبيل المثال، من المرجح أن تكون النظم الخاصة والشبكات الداخلية الخاضعة للسيطرة المحلية أو لسيطرة أرباب العمل الخاصين بك (مع مراعاة الأجهزة، والبرمجيات، وسلطات منح الشهادة، وإدارة نظام التجميع والشبكات) أكثر أهلية للثقة.

الأمر الجديد الذي يشكّل مصدراً لمخاطر جديدة هو حجم بعض السحابات وقابليتها للتوزيع. فالسحابات هي نظم موزعة كبيرة تخضع لسيطرة مركزية محدودة. تُعتبر السحابات التي تتيح النفاذ إلى كميات هائلة من المعلومات (مثل Google و Amazon) موارد قيمة جداً، ومع ذلك، يمكن للسحابات الأخرى التي تخزن المعطيات الخاصة بك (جنباً إلى جنب مع معطيات الآخرين) أن تطرح مشاكل خطيرة تتعلق بالوثوق بكيانات يُحتمل أن تكون غير جديرة بالثقة.

إن إطلاق اسم "سحابة" على المفهوم القديم للنظم الكبيرة المشتركة الموزعة قد يكون مُضللًا، فهو يُطلق مصطلحاً رناناً جديداً، ويخفي من جهة ثانية المشاكل الكامنة وراء المخاطر التي اصطدم بها المصممون والمديرون لسنوات.

وقد تجاهل بعض مقدمي الخدمات السحابية العديد من المخاطر القديمة، ومن الجلي أنهم غافلون لدرجة كبيرة عن المخاطر الجديدة كذلك. من الواضح أن الحوسبة السحابية هي ببساطة الحوسبة عن بُعد التي شكّلت أحد الأسباب الرئيسية لإنشاء ARPANET - للسماح للأشخاص الموجودين في منطقة زمنية ما بالاستفادة من الموارد غير المستعملة في المناطق الزمنية الأخرى خلال ساعات معينة من اليوم. وقد أصبح هذا بكل وضوح ذا فائدة أكبر مع ظهور الإنترنت، نظراً لشمولها جميع أنحاء العالم. وبالمثل، فإن التخزين السحابي هو ببساطة تخزين عن بُعد، شاع استعماله في البداية كنسخة احتياطية خارج الموقع لأسباب واضحة تتمثل في التسامح مع الخطأ، والتأهب لحالات الطوارئ، وغيرها من الأسباب.

يتمثل أحد مخاطر تحديد تخزين بعيد للنسخة الاحتياطية خارج الموقع كـ "تخزين سحابي" في كون هذا المصطلح يحجب وجود بدائل داخلية، مثل الممارسة الشائعة المتمثلة في التسجيل الدوري للقطات لنظام الملفات على وسيط صغير منفصل، والاحتفاظ بها في مكان آمن. يمكن أن يكون لهذا الأمر أهمية خاصة بعد هجمات الاختراق المؤدية التي قد تؤثر سلباً في النظام وذلك بإدراج برامج ضارة، وبرامج تُلصص، وهلم جزءاً. زيادةً على ذلك، الأرشفة عن بُعد - خاصة إذا

توزعت على نطاق واسع بين مستودعات مختلفة - تجعل المستخدمين غير متحققين من أن معلوماتهم ما تزال قابلة للاسترجاع بشكلها الأصلي (ما لم يكونوا قد استرجعوها بالفعل).

وقد نُقِلَ عن Ron Rivest قوله: "تبدو الحوسبة السحابية لطيفة جداً ورائعة وآمنة ... ينبغي لنا فقط الانتباه إلى المصطلحات إن كنا ندعوها حوسبة المستنقع. أعتقد أنك قد تمتلك التفكير الصائب".²

لإعادة صياغة اقتباس غالباً ما يُنسب إلى Roger Needham، أو Butler Lampson، أو Jim Morris، إذا كنت تعتقد أن الحوسبة والتخزين السحابيين هما الحل لمشاكلك، فأنت لا تدرك تلك الحلول الممكنة، ولا تفهم مشاكلك.

تجدر الإشارة هنا إلى أمثلة قليلة على الاستعمالات الحديثة ذات الصلة والمحفوفة بالمخاطر. (يمكن الاطلاع على مزيد من المعلومات الأساسية عن البنود التسعة الأولى وأمثلة أخرى على التسويات السحابية في منتدى المخاطر في ACM: <http://www.risks.org>).

- تعرضت خدمات المشاركة المقدمة من Dropbox للاختراق، ونجم ذلك عن ثغرة أمنية في مخطط مشاركة الروابط الذي يتبناه. وقد قام الجناة أيضاً بنشر الاستغالات.
- انتزعت مايكروسوفت 22 نطاقاً من نطاقات NO-IP الشائعة الاستعمال، فيما يبدو أنه تخطت لأمر المحكمة الاتحادية.
- انقطعت خدمات الوب من Amazon (مدة قصيرة) عدة مرات. فقد قام مهاجم بتدمير Code Spaces (وهو مستودع قيم للرماز المصدري أنشئ باستخدام مرافق AWS الخاصة بـ Amazon) تدميراً مؤثراً طلباً لفدية.
- امتلكت Cisco Systems مفتاح تعمية مضمناً في مدير VoIP الخاص بها، أتاح السيطرة غير المخوّل بها على المعدات الحساسة للتراسل.
- عملت Cryptolocker وبرامج أخرى لانتزاع الفدية على تعمية المعلومات المخزنة بشكل قسري، وطالبت بأموال لفك تعميته (مع أنها لم تقم في بعض الحالات بفك التعمية حتى بعد تسلمها للفدية!). ومع أن معظم هذه الهجمات قد استهدفت مستخدمين أفراداً، فإنّ من الواضح أن فرصة شن هجمات على مستودعات التخزين البعيد تشكل خطراً. (وكما ورد، فقد تم في الآونة الأخيرة إنشاء موقع وب مضاد للهجمات).
- أوقف العمل بـ TrueCrypt (لتعمية كامل القرص) كبرمجية متاحة المصدر من قبل واضعيها الذين يحملون أسماء مستعارة، نظراً لأنها قد تحتوي على مسائل أمنية غير محلولة. (ما يزال هناك شك في خطورة وتأثير هذه المسائل المحتملة، وفي إمكان إيجاد طريقها إلى مصدر البرنامج).
- حدثت أمور مشابهة لـ Lavabit التي أتاحت ميزات الخصوصية والأمان في خدمات البريد الإلكتروني لأكثر من 400,000 زبون، وقد تم سحبها بعد مضايقات قانونية مطوّلة حاولت فرض إرساء معدات للمراقبة.
- انتزعت Megaupload.com من قبل السلطات التي حظرت المستخدمين الشرعيين وغير الشرعيين.
- أفلست Nirvanix مالياً، وأعطى مستخدموها مهلة أسبوعين للمغادرة.
- كانت المحادثات المختلفة التي جرت في Black Hat و DEF CON في آب مخيبة للأمال لبعض الشيء. باختصار، بدا كل جهاز معرضاً من حيث الجوهر للخطر لامتلاكه كلمة سر رئيسية ثابتة مضمّنة في النظام، إضافة إلى وجود الكثير من نقاط الضعف الخفية أيضاً. تُعتبر هذه الأخبار قديمة للبعض، لكنها قد تكون صادمة لآخرين.

ومن بين المخاطر القديمة التي ما تزال منتشرةً الاستعمال الداخلي للتخزين المحلي الذي يمكن أن يؤدي إلى عطب في العتاد، وإلى فشل في برمجيات قواعد المعطيات. قد تنتهي جميع النسخ المكررة في الواقع في مستودع سحابي واحد معرض للهجوم. فضلاً على ذلك، قد يكون من غير الممكن الاستمرار في دعم مصاعغات المعطيات القديمة. ما يزال التخزين المحلي يتطلب الانتباه إلى النسخ الاحتياطي الذي يمكن استرجاعه بنجاح - بعد مرور سنوات عديدة في بعض الحالات. مع ذلك، إذا تمت تعمية المعلومات الأصلية، تصبح القدرة على إدارة واستعادة المفاتيح القديمة مسألة حرجة. تظهر مخططات التعمية الفعالة على نحوٍ متزايد في الآونة الأخيرة ضمن مجتمعات البحوث لإثبات حيابة المعطيات، وإثبات قابلية استرجاع المعطيات. ولسوء الحظ، فإن التقنيات البسيطة وغير المكلفة من هذا النمط لم تجد إلى الآن طريقها للانتقال من المبادئ النظرية إلى الممارسة. ربما تكون الجهود التي يبذلها مقدمو السحابة لضمان قابلية استرداد المعطيات المخزنة الخاصة بهم ذات فائدة أكبر. قد يكون جيداً أنه، في المتوسط، تُدار أنظمة مقدمي السحابة إدارة أفضل من مجموعات تقانات المعلومات التابعة للعديد من المؤسسات والوكالات. على الأقل، من المؤكد أن مستخدمي السحابة يأملون هذا! ومع ذلك، لا تزال هناك مخاطر مختلفة.

تتطوي إحدى المشاكل القديمة المتفاقمة على إمكان حذف المعلومات المنتشرة في كل مكان. إن وجود نسخ منتشرة وإصدارات مختلفة يُضخّم تضخيماً واضحاً نتيجة النسخ التي تم تكرارها بغية استرجاعها. وقد انتشرت المواقع المتطابقة على الإنترنت إلى حد يفوق قدرة أي شخص على تتبع الإصدارات المخفية. مع القيام بتخزين المعلومات في أحد المستودعات البعيدة غير الخاضعة للمساءلة، سيبدو دوماً الحذف المنتشر على أنه أمر مشكوك فيه. إلى جانب ذلك، فإن التوجهات التي يمكن أن تتجح في الحذف المنتشر ربما تكون أيضاً ضحية للحذف غير المقصود أو الخبيث. في هذه الحالة، قد يكون من المرغوب فيه وجود آلة للزمن.

تحتاج العديد من المخاطر ذات الصلة الاجتماعية أيضاً إلى الدراسة، مثل تلك الإصدارات المختلفة للمعطيات المزيفة، ووجود معلومات في إصدارات للبحث غير مطابقة تماماً للأصل تدّعي أنها المعلومات نفسها، والحالات التي يرغب فيها الأشخاص أو المنظمات بإخفاء بعض المعلومات إخفاءً كاملاً.

وبطبيعة الحال، تعرض القوانين والتشريعات الدولية أيضاً العديد من المشاكل - أولاً بسبب غموضها أو توسّعها المفرط، وثانياً بسبب عدم اليقين المحيط بأصول ووجهات المعطيات وطلبات الموارد الأخرى. على سبيل المثال، إذا أصرت أمة على ضرورة تخزين جميع المعلومات العائدة لمواطنيها في نظم تخضع لإشرافها القانوني الخاص بها، فكيف يمكن ضمان ذلك عندما يكون من السهل جداً تخريبها، وعندما تكون ملكيتها في حد ذاتها غامضة؟ إضافةً إلى ذلك، يجب أن نكون مدركين لمخاطر المراقبة المنتشرة في كل مكان في الموارد البعيدة غير الخاضعة للمساءلة، والتي تكون في بعض الحالات غير معروفة.

من المحتمل أن يكون كل حاسوب أو كيان بشري تقريباً غير جدير بالثقة، بالنظر إلى الحوادث وسوء الاستعمال المتعمد، والهجمات. يتمثل أحد الأمثلة التي لا تزال مصدراً للمشاكل في الاستعانة بمصادر خارجية في الانتخابات، مع مراعاة الاعتماد على النظم والبرمجيات الاحتكارية، وموارد الحوسبة وقواعد معطيات التسجيل، والشبكات (سواء المفتوحة منها أو الخاصة)، وقبل كل ذلك الاعتماد على الأشخاص المشكوك في أهليتهم للثقة، ويوضح هذا المثال جيداً طبيعة المخاطر من البداية إلى النهاية انطلاقاً من بداية الحملة الانتخابية، ووصولاً إلى النزاعات التي تنتج عن مصادر الخطأ، والغش، والارتباك - مع ما يصاحب ذلك من الخوف، وعدم اليقين، والشك. يُعدُّ سوء الاستعمال الداخلي أمراً خطيراً في

جميع الموارد المشتركة، ويكتسب أهمية خاصة في الانتخابات (لوحظت حالات كثيرة في منتدى المخاطر Risks Forum وغيرها). في هذا المثال، تمثل الاستعانة بمصادر خارجية من الكيانات غير الخاضعة للمساءلة إشكالية كبيرة. على الرغم من المخاطر التي ناقشناها هنا، هناك بعض الآمال بإيجاد بدائل ببناءة. تنتهج المجتمعات البحثية توجهات مختلفة نحو الأجزاء المكونة لهذا اللغز، ولكنها نادراً ما تمتلك توجهات نحو النظم ككل. ونتيجة لذلك، يتعلق العديد من مقالات سابقة حول موضوع هذا البحث باستعمال أو إساءة استعمال الموارد عن بُعد - حتى وإن كانت تركز على المشاكل التي كانت تعتبر سابقاً محلية. على سبيل المثال، التعمية التي تُدار بشكل منفرد من قبل المستخدمين النهائيين للمعلومات المخزنة عن بُعد بمصاغات معمّاة غالباً ما توصف بأنها حل لمشكلة الحاجة إلى الوثوق بمقدّم غير موثوق به للتخزين البعيد. تستطيع التعمية المتماثلة الشكل (homomorphic cryptography) إمكان السماح بإجراء الحسابات على المعلومات المعمّاة، دون الحاجة لفك تعمية هذه المعلومات. يمكن لهذه التوجهات زيادة سرية المعلومات، كما يمكنها إتاحة وسيلة لمشاركة المعلومات بواسطة مفاتيح تعمية مشتركة خارج النطاق. ومع ذلك، تبقى المعلومات عرضة لتسويات أخرى مثل الحذف غير المقصود أو الحذف الخبيث، وهفوات العقود مع مقدمي الخدمة عن بُعد، وفقدان مفاتيح التعمية، وعدم توفر الخدمات، ومراقبة الاستعمال العدائي، وهلمّ جرّاً. وكما هو صحيح بوجه عام، تصبح إدارة المفاتيح في حد ذاتها خطراً أساسياً. وزيادةً على ذلك، تبقى المخططات الملائمة لاستعادة المفاتيح المفقودة (برمجيات الباب الخلفي على سبيل المثال) عرضة دوماً لإساءة الاستعمال - كما هو الحال في الأبواب الخلفية التي يمكن أن يُساء استعمالها من قبل الأشخاص من الداخل أو عن طريق الهجمات الخارجية.

وقد درس Virgil Gligor¹ بعض المخاطر الكامنة في المحاكاة الافتراضية ضمن سياق مماثل جداً لما يجري فحصه في هذا المقال. في المحاكاة الافتراضية بعض الجوانب المشتركة مع التجريدات التي يتيحها التنفيذ عن بُعد والوصول عن بُعد، بمعنى أن هناك واجهات محددة جيداً للتعامل مع كلا الحالتين - سواء أكانت بعيدة افتراضياً أم مكانياً. هناك أيضاً أسئلة حول صدقية الآليات الكامنة وراء فرض التجريدات الافتراضية - على سبيل المثال، الكبسلة، والنفاذي، أو من ناحية أخرى إخفاء الضعف في الطبقة الدنيا. يتناول مقال Gligor تناولاً صريحاً بعض المواضيع المشار إليها هنا، ويستحق قراءة متأنية من قبل أولئك القراء الذين يرغبون في الحصول على معلومات أساسية أعمق مما هو مُدرج هنا. أوكد أن للسحابات القدرة على تقديم فوائد حقيقية وملموسة. كما أنها تجلب العديد من المخاطر التي يمكن أن تكون مختبتة وراء بساطة تجريد السحابة. يجب أن تفكر ملياً في هذه المخاطر عند تصميم، واختيار، وتشكيل الخدمات السحابية الخاصة بك.

المراجع

- [1] Gligor, V. Security limitations of virtualization and how to overcome them. Security Protocols Workshop, SPW 2010, Cambridge, U.K., 2010.
- [2] McMillan, R. Cloud computing a security nightmare, says Cisco CEO. Computerworld; http://www.computerworld.com/s/article/9131998/Cloud_computing_a_security_nightmare_says_Cisco_CEO.

بيتر نيومان (neumann@csl.sri.com) هو كبير العلماء الرئيسيين في مختبر علوم الحاسوب في SRI الدولية، ورئيس جلسة في ACM Risks Forum.

يسبر الباحثون الأمن بواسطة التعميم

RESEARCHERS PROBE SECURITY THROUGH OBSCURITY*

Chris Edwards

ترجمة: د. أميمة الدكاك

مراجعة: د. أديب بطح

يحمي التشويش الرماز بجعله عصبياً على الاختراق، بحيث لا يساعد النفاذ إليه القرصنة على فهم كيفية عمله.

قادت محاولة بناء طريقة تمنع الهندسة العكسية للبرمجيات إلى كم هائل من الأفكار التي يمكن أن تعيد تشكيل عالم التعمية، ولو أنّ الهدف الأساسي الذي يكمن في إخفاء كيفية عمل البرمجيات يمكن أن يُثبت أنه صعب الإنجاز من الناحية العملية.

قد ترغب الشركات في حماية الملكية الفكرية المخزنة في برمجياتها بواسطة إخفاء الطريقة التي تعمل بها هذه البرمجيات. يكمن أحد الخيارات في استعمال التعمية لبعثرة الرماز المخزن في الذاكر الدائمة (غير الطيارة)¹، بحيث لا يظهر الرماز نصاً واضحاً إلا بعد أن يجري فك تعميته ويصبح جاهزاً للتنفيذ. لبعض المعالجات المُضمَّنة حالياً محركات فك تعمية على الجذاذة (الدائرة) نفسها، لتجنّب ظهور تعليمات البرنامج الفعلية على مسرى الذاكرة، بحيث لا يمكن لقرصان مزود بمحلل منطقي² أن يعترض تعليمات الآلة. ومع ذلك، مازال من المحتمل أن يتمكن المهاجمون من استعمال تقنيات مثل تحليل "قناة جانبية"³ لاستعادة مفتاح التعمية الفعلي واستعماله.

يعتمد التعميم منهجاً آخر، بجعل الرماز، أساساً، مضللاً وغير قابل للاختراق، بحيث لا يساعد النفاذ إليه المهاجمين على فهم كيفية عمله. يكمن النهج الذي يتبعه الباحثون في تعريف (تحديد) التعميم المفيد في تطبيق فكرة أن يكون البرنامج غير مفهوم، بحيث لا يكون بإمكان العدو استخلاص المزيد من المعلومات من البرنامج المُعتمَّ أكثر مما لو كان ينفذ فقط إلى تجزير صندوق أسود.

مفهوم الصندوق الأسود الافتراضي (VBB) Virtual Black Box هو مفهوم طوره بواز باراك Boaz Barak في مقالة علمية كتبها في عام 2001، ثم تابع تطويره معهد وايزمان للعلوم Weismann Institute of Science وعاملون معهم من مجموعة جامعات أمريكية، وهم من وضعوا أساس العمل الحالي ولكن يبدو أنه أيضاً قد حطّم آمالهم. يبدو أن

* تُشير هذا البحث في مجلة *Communications of the ACM*، المجلد 57، العدد 8، آب (أغسطس) 2014، الصفحات 11 – 13.

¹ الذاكر التي تحتفظ بالمعلومات وإن قُطعت عنها التغذية الكهربائية. (المترجم)

² أداة إلكترونية تلتقط الإشارات من جهاز رقمي أو دائرة رقمية وتظهرها على هيئة مخطط توقيت أو مخطط حالة أو لغة تجميع، وغيرها. (المترجم)

³ طريقة للهجوم على نظم التعمية يعتمد على المعلومات المأخوذة من تجزيه المادي وليس البرمجي أو البحث الشامل. (المترجم)

وصفت الباحثة بعد الدكتوراه post-doctoral إيليت بويل Elette Boyle من جامعة كورنيل- التي قدمت عملاً ذا صلة بنموذج التعميم في مؤتمر نظرية التعمية cryptography في سان دييغو- هذه المقالة "بالمقالة الرُّشِمية"⁴ التي فتحت أبواب بحوث التعميم على مصراعيها ثانية. وقالت: "بعد ذلك العمل، كان هناك سيل من النتائج الجديدة، ولكن يصعب العمل عليها. عليك الحصول على دارات مكافئة تماماً لهذه النتائج، بحيث تعطي نفس الخرج إذا أُعطيت الدخل نفسه". مع أن التعميم غير المميز IO يبدو محدوداً من حيث القابلية للتطبيق، لأنه فقط يجعل من المستحيل على المهاجمين إيجاد أي فرق له معنى بين نسختين معتمتين لنفس البرنامج، فإن وجود هذه التقنية القابلة للاستعمال قد برهن من زاوية هجوم جديدة استحالة البرهان. القضية هي البناء على هذا الأساس لإنشاء نظرية تعميم متعدد الاستعمال. حتى في حالة IO الذي طوره بدايةً جريج وزملاؤه، كانت أنماط البرامج التي يمكن التعامل معها محدودة جداً. وقد افترض البرهان ألا يُسَمَح للمهاجمين إلا بمجموعة محددة من الأفعال (الإجراءات) وهذا ما لا يمكن إجبار المهاجمين على فعله في العالم الحقيقي.

خلال زمن أقل من عام، وسَّع الباحثون مجال البرامج التي يمكن لـ IO ونماذج التعميم الأخرى ذات الصلة مهاجمتها، وخفضوا كثيراً الحاجة إلى افتراضات صناعية. ومع ذلك، يقول جارج سنستغرق أكثر من عقد من الزمن لنصل إلى النقطة التي يتوسَّع فيها التعميم إلى أبعد من أشكال IO بحيث نتمكن من إخفاء أعمال البرمجيات. ويتابع القول: سيكون من الصعب الحفاظ على سرعة التطور الذي حصل خلال العام الماضي.

يقول جارج: نحتاج إلى عدة فُتُوح ولكن لسوء الحظ تحدث الفتح في لحظات عشوائية، ومن الصعب التنبؤ بوقت حدوثها. لقد حدث الكثير من الفعاليات منذ صيف 2013 ولكن للأمانة، هي خُرُج الفورة الذهبية التي حصلت آنفاً. مع أن ظهور التعميم المميز IO لم يمكِّن من الحصول على التعميم المتعدد الاستعمال فإنه اقترح إمكانيات أخرى للباحثين في حقل التعمية.

يقول ووترز Waters: "حين طُرِح تعريف التعميم أول مرة، وكذلك في مقالة باراك وزملائه أيضاً، لم يحظَ بكثيرٍ من الاهتمام، نسبياً، لأن كيفية استعماله لم تكن واضحة. ومع ذلك، بيّن عمل حديث أن هذا التعميم أداة هامة جداً لتصميم التعمية".

بيّن باحثون أن IO يمكن أن تكوّن دعامات (أساسات) لتقانات جديدة للتعمية. أحد التطبيقات الممكنة هو "التعمية الوظيفية" التي تمكِّن من إخفاء أجزاء من نفس كتلة المعطيات انتقائياً من عدة مستخدمين باستعمال مفاتيح فك تعمية متعددة، عوضاً عما يجري في الحالة الحالية، التي يُتاح فيها خيار تعمية وحيد بتوفير مفتاح وحيد يفك تعمية كامل الملف. يقول وترز: "ثمة فجوة وظيفية بين كيف نفكر في اقتسام المعطيات والآليات الفعلية للقيام بذلك"، ويضيف أن ما لدينا اليوم هو حلول هندسية تحاول التَّغَلُّب على المشكلة: "هذا ما أتى عليه علم التعمية: إعادة التفكير فيما نعنيه حقيقةً بالتعمية".

بيّن باحثون مثل ووترز أن استعمال التعمية المبنية على إطار عمل يقترحه IO سيمكِّننا من الحصول على أنماط تعمية أكثر مرونة. على سبيل المثال، التعمية الوظيفية المعتمدة على IO تمكِّننا نظرياً من توفير مفاتيح جديدة للمستخدمين يمكنها فتح مقتطعات مختلفة من المعطيات حتى بعد تعمية المعطيات بمفتاح رئيسي master key. واقترح بويل وباحثون آخرون أنواعاً أخرى من التعمية مستخلصة من IO تسمح بفتح المعطيات باستعمال أنواع أخرى من

⁴ أي التي تشمل على بذور التطور في المستقبل. (المترجم)

المعطيات مثل حل مسألة رياضية بصعوبة NP⁵. ومع أن هذه الأنواع من التعمية التي بدأت بالظهور قد لا يكون لها تطبيقات مباشرة، فإن الباحثين واثقون أن هذا العمل سيوسّع كثيراً قابلية تطبيق التعمية. والأمل النهائي هو أن العمل في هذا الحقل سيجنبنا الحاجة إلى استعمال وظائف متخصصة وليس فقط توفير مستوى من التعقيم بل إلى بناء أمن أقوى مما لدينا الآن (اليوم). يُحتمل ألا تعتمد هذه الطرائق كثيراً على التعقيد الحسابي. ستكون النتيجة ألا تُضعف الحواسيب، التي يرتفع أداؤها باستمرار، التعمية التي نستعملها. وإذا كانت الأبحاث الأخرى تحطم الأمل في بناء معتم برمجيات متعدد الاستعمالات، ومُبرهن رياضياً، فقد برزت بعض الأشكال القابلة للاستعمال التي يمكن أن تمكّن من إخفاء عناصر مفتاحية من البرامج، وتسليم تعقيم جزئي على الأقل، على أمل أن يكون هذا التعقيم قابلاً للبرهنة رياضياً.

ويقول عمر بانيث Omer Paneth وهو طالب دكتوراه في قسم علوم الحاسوب في جامعة بوسطن والباحث في التعمية: "تخيّل بعض البرمجيات التي تحتاج إلى تلفية، وأنها تنهار في حال استعمال إدخلالات سيئة (غير مناسبة). يمكنك استعمال بقعة برمجيات patch ترشح هذه الدخول السيئة وتعتم عليها بحيث لا يراها الناس ولا يتعلمون الدخول السيئة التي يمكن أن تجعل البرمجيات تنهار". ويقول بانيث إن هذا النشاط ممكن إذا كان المرشح أحد عناصر طائفة التوابع المراوغة "evasive functions" التي تعيد القيمة صفر لكل الدخول تقريباً باستثناء نتائج مفيدة لبعض الدخول المختارة. وبحسب بانيث، فإن التوابع المراوغة evasive تصعب على العدو أن يتعلم أي شيء عن سلوكها، وهي تنجو أيضاً من برهان الاستحالة الذي توفره مقالة العام 2001. ومع ذلك، يوقن باحثون مثل جارج أن العلم سيتوسع أكثر ليسمح بانتشار أوسع لاستعمال التعقيم.

ويقول جارج: "سأمتنع عن اقتراح أنه سيكون بإمكاننا قريباً جداً تنفيذ برمجيات مايكروسوفت المكتبية Microsoft Office بشكل معتم، ولكنني لا أدرك لماذا لا يمكننا تعقيم برامج من حجم معقول. هذا يحتاج إلى أبحاث ضخمة، إضافة إلى جهود هندسية ضخمة، ولكنني لا أدرك لماذا لا يمكننا القيام بها".

قراءات للاستزادة

- Barak. B., Galdreich, O ., Impagliazzo, R ., Rudich. S., Sahai, A, Vadhan, S ., Yang, K. "On the (Im) possibility of Obfuscating Programs, Advances in Cryptology (2001) 2010 revision: <http://www.wisdom.weizmann.ac.il/~oded/PS/obf4.pdf>
- Garg, S., Gentry. C ., Halevi, S., Raykova, M., Sahai A., Waters B. Candidate Indistinguishability Obfuscation and Functional Encryption for all circuits. 54" Annual Symposium on Foundations of Computer Science (FOCS), (2013).
- Sahai, A., Waters, B. "How to Use Indistinguishability Obfuscation: Deniable Encryption, and More", IACR Cryptology ePrint Archive 2013: 454 (2013). <http://eprint.iacr.org/2013/454>.
- Barak, B., Bitansky, N., Canetti, R ..Kalai, Y.T., Paneth, O., sahai, A. "Obfuscation for Evasive Functions", Theory of Cryptography, Lecture Notes in Computer Science, Volume 8349, pp26-51 (2014).
- Gentry, C ..Lewko, A., Sahai, A., Waters, B, "Indistinguishability Obfuscation from the Multilinear Subgroup Elimination Assumption", IACR Cryptology ePrint Archive 2014: 309 (2014) <http://eprint.iacr.org/2014/309>.

⁵ NP، Non Polynomial أي لا يمكن حلها بزمن كثير حدودي. (المترجم)

أنظمة إسقاط المصوّرات التجسيمية توفر حياةً أبدية

HOLOGRAPHIC PROJECTION SYSTEMS PROVIDE ETERNAL LIFE*

Esther Shein

ترجمة: د. محمد سعيد دسوقي
مراجعة: د. محمد عباسي

تساعد الحيل الضوئية الفنّانين الموتى على الاستمرار بأدائهم.

عندما توفي الممثل فيليب سيمور هوفمان Philip Seymour Hoffman على نحوٍ غير متوقع في شباط عام 2014، كان قد بقي له أسبوع على الأقل من التصوير في دور الطائر Mockingjay، الجزأين الثالث والرابع من سلسلة الأفلام Hunger Games التي لاقت نجاحاً كبيراً. من حسن الحظ، أنّ هوفمان شأنه شأن فنّانين آخرين متوفّين من قبله، ومنهم مغني "الراب" توباك شاكور Tupac Shakur والمغنون مايكل جاكسون Michael Jackson، وفرانك سيناترا Frank Sinatra، وإلفيس بريسلي Presley Elvis، كان لديه القدرة على مواصلة العيش وسحر جمهور جديد باستعمال ما يسمى أنظمة إسقاط المصوّرات التجسيمية 3D holographic projection systems.

التصوير التجسيمي holography هو تقنية لتسجيل صدر الموجة لمشهد ثلاثي الأبعاد على صورة ذات بعدين، وفقاً لتينغ تشونغ بون Ting-Chung Poon في عام 2013، وهو أستاذ في قسم برادلي Bradley للهندسة الكهربائية وهندسة الحاسوب في المعهد التقني في ولاية فرجينيا. فقد كتب بون وزميله تسانغ P.W.M. Tsang في بحثهما بعنوان مراجعة نظرية وتطبيقات إطار عمل تسجيل صدر الموجة في توليد ومعالجة المصوّرات التجسيمية الرقمية Review on Theory and Applications of Wavefront Recording Plane Framework in Generation and Processing of Digital Holograms: "قبل بضعة عقود، كان يمكن إدارة هذه الإجراءات باستعمال الوسائل البصرية فقط، بمزج موجة الغرض المنظور بحزمة مرجعية، وتسجيل الصورة الناتجة على فيلم تصويري. ومع التقدم السريع لتقنيات الحوسبة في السنوات الأخيرة، أمكن محاكاة آلية تشكيل المصوّرات التجسيمية الضوئية بالوسائل الرقمية."

* تُشير هذا البحث في مجلة Communications of the ACM، المجلد 57، العدد 7، تموز (يوليه) 2014، الصفحات 19 – 21.

دور الحواسيب

تحسب هذه الطريقة، المتعارفة بالتصوير التجسيمي المولد بالحاسوب (Computer Generated Holography (CGH، صورة الانعراج الناشئة عن غرض ثلاثي الأبعاد، وتضيف موجة مرجعية لإنتاج مصوّر تجسيمي رقمي يمكن عرضه على جهاز الكتروني، كما يقول بون، وهو أيضاً زميل في الجمعية الضوئية (Optical Society). "يتناول التصوير التجسيمي المولّد بالحاسوب طرائق توليد المصوّرات التجسيمية رقمياً. يمكن طباعة المصوّرات التجسيمية بالتتالي على فيلم أو تحميلها على معدّل ضوئي مكاني (Spatial Light Modulator (SLM لإعادة تركيب تصوير تجسيمي"، كما يشرح بون. يبيّن بون، أنّ جملة "التوليد الرقمي للمصوّرات التجسيمية" تشير إلى الحسابات التي يقوم بها الحاسوب لبناء مصوّر تجسيمي لغرض ثلاثي الأبعاد. أصبح التصوير التجسيمي الرقمي (Digital Holography (DH ممكناً بسبب التقدم المستمر في الحوسبة وفي المسح الضوئي، إضافةً إلى إتاحة تخزين رقمي بسعة كبيرة وتقانات اتصال عريض الحزمة.

دمج تقانات القرن التاسع عشر في تقانات اليوم

كان التصوير التجسيمي نعمة على صناعات التسالي والفنون خصوصاً، مع تقانة يمكن استعمالها لإعادة ممثلين متوفين

"إلى الحياة مجدداً". على سبيل المثال،

استعملت شركة موزيون Musion 3D، ومقرها لندن، والتي تطور أشكال مجسّمات ثلاثية الأبعاد وتسوقها وتنتجها وتبيّنها، تقانة Eyeliner لإحياء عدد من الفنانين المتوفّين، منهم الممثل ليس داوسون Les Dawson للظهور على محطة ITV في حزيران 2013 - في حدث كان مخطط له قبل وفاته في حزيران 1993.

أنشئت شركة Musion Eyeliner في عام 1995 عندما كُفّ مؤسس الشركة أو ماس Uwe Maass وهو مهندس عروض ليزرية، بإنتاج عرض شبح الفلفل Pepper's Ghost display لمصلحة مصنع الكريستال سواروفسكي Swarovski. يدلّ تعبير شبح الفلفل على نظام إسقاط ضوئي ظهر في القرن التاسع عشر، ينتج



مصوّر تجسيمي لمايكل جاكسون وهو يؤدي دوراً على خشبة المسرح خلال حفل توزيع جوائز الموسيقى بيلبورد Billboard في لاس فيغاس 2014، بعد مُضيّ ما يقارب خمس سنوات على وفاة نجم "البوب".

مؤثرات مسرحية شبه حية باستعمال زجاج مسطّح أو أفلام شفافة وإضاءة خاصة.

يقول يان أو كونيل Ian O'Connell مدير Musion 3D "تطلّب عرض سواروفسكي إيجاد بديل عن استعمال صفائح الزجاج التي كانت حتى ذلك الوقت المادة المستعملة في خلق الوهم illusion، كما في منزل ديزني المسكون Disney's Haunted House في كاليفورنيا". وبعد البحث عن مركبات بوليمرية متعددة، وجد ماس شريحة تصوير عريضة من شركة هويشت Hoechst (ميتسوبيشي Mitsubishi حالياً)، التي وفّرت المستوى المطلوب للصورة. هذا الإنجاز الكبير الذي أمكن تحقيقه، إضافةً إلى التطور الحاصل في الإسقاط الفيديوي العالي الدقة خلال هذه المدة، وفراً إعادة إحياء منصة شبح الفلفل."

يقول أو كونيل: ثمة ثلاثة مكونات مفتاحية لنظام Eyeliner: رقائق مطوّرة خصوصاً لعكس الصور الصادرة عن أجهزة الإسقاط الفيديوية العالية الدقة، والمنصة، والمحتوى. تُشكّل المنصة بأسلوب مشابه لمجموعة موسيقية حية مع صوت وإضاءة. وتُشكّل الرقائق بزوايا 45 درجة على كامل المنصة، "مشدودة ضمن إطار بحيث لا يراها الجمهور المشاهد". يكون المحتوى عادةً صوراً ثلاثية الأبعاد مولدة بالحاسوب (3D Computer-Generated Imagery (CGI)، ممزوجة بفيلم 35mm أو فيديو عالي الدقة (1080p or 1080i) لإتمام التأثير. ويقول أيضاً "تُستعمل الحواسيب في إنشاء المحتوى، وفي أتمتة عمليات المنصة مثل مزامنة التحكم الضوئي مع الفيديو، وأخيراً في إعادة تشغيل محتوى الفيديو". يُستعمل Eyeliner أساساً في أنشطة عرض المنتجات مثل عروض التجارة، وعروض الأزياء، والحفلات الموسيقية، بحسب "أو كونيل". ويضيف "مع ذلك، فإن استعمالات جديدة يجري إنشاؤها في صناعات مختلفة". على سبيل المثال، في شباط، كشفت شركة إنتربلوك Interblock S. A. التي تورّد منتجات ألعاب إلى نوادي القمار casinos، النقاب عن قاعة التصوير التجسيمي الخاصة بها باستعمال مولد الأرقام الحاسوبي الخاص بها مرفقاً بتقانة Musion في معرض نوادي القمار الدولي لعام 2014 International Casino Exposition/ICE 2014.

أنشأت شركة AV Concepts، وهي مزوّدة لتقانة التصوير التجسيمي الغامرة، عروضَ تصوير تجسيمي باستعمال رقائق تصوير تجسيمي خاصة ومؤثرات ضوئية ومخدم وسائط خاص "للعرض التصويري المتغير بسهولة" Liquid Scenic projection. يشرح مهندس التصميم الإبداعي ألوك وادوانير Alok Wadhwanir "يأخذ المخدم صوراً غير مضغوطة نهائياً ويستعمل أجهزة إسقاط عالية السطوح متعددة تتكامل بسلاسة لإنشاء عروض قوية شبه حية". تُستعمل رقائق مشدودة بقوة لعكس الصور.

يقول وادوانير أنجزت شركة Tempe عدداً من المشاريع باستعمال تقانة عرض التصوير التجسيمي في صناعات متنوعة، تتضمن العمل مع سلسلة فنادق كبيرة لإحضار مديرتها التنفيذي إلى المنصة على نحوٍ افتراضي، مع بائع أحذية لإطلاق حذاء رياضي جديد وعرض تقانته الداخلية على الجمهور باستعمال التصوير التجسيمي، ومع الفنان دولي بارتون Dolly Parton لإحضار أشباح عيد الميلاد المصوّرة تجسيمياً إلى موضوع المنتزه الخاص بها Dollywood theme park's rendition of A Christmas Carol.

يحدّر ميكائيل جود Michael Jude، مدير برنامج خدمات اتصالات المستهلكين في مؤسسة البحث Frost & Sullivan، من أنّ ما يفعله Musion وآخرون ليس إسقاط تصوير تجسيمي صرفاً. ويتابع، ليس ثمة أنظمة إسقاط تصوير تجسيمي فعلية متاحة تجارياً، مضيفاً أنّ ما تقوم به هذه الشركات هو ببساطة إنشاء تأثيرات خاصة على طريقة شبح

الفلفل. "إنّه عرض ضوئي معروف جيداً مرفق بصور مولدة حاسوبياً ... يتطلب الأمر استعمال الليزر لتوليد مصوّرات تجسيمية حقيقية."

يقول جود: حالياً، الطريقة الوحيدة لعمل تصويرات تجسيمية هي بالنظر خلال صفيحة/صورة ثلاثية الأبعاد تشكل نمط التداخل. "إذا أردت أن تعيد إلفيس Elvis، يجب أن تشكّله في مكان ما قبل إسقاط [الصورة] ... وهذا يتعلق تعلقاً شديداً بالحاسوب، ولا يكون بالزمن الحقيقي."

يتطلب توليد مصوّر تجسيمي ارتداد الليزر عمّاً تريد النقاطه. "قأنت تضيء ذلك [الغرض] بحزمة ليزرية ويرتد ذلك الضوء عن [الغرض] ويضرب صفيحة تصوير ضوئي،" يشرح جود. "ثم تسلط حزمة الليزر نفسها على الشيء في ذلك الوقت الذي تضيء فيه صفيحة التصوير الضوئي، لإنشاء التداخل." بمجرد إنشاء نمط التداخل المسطح على الصفيحة، يمكن للمستعمل تسليط الليزر خلالها لتوليد مصوّر تجسيمي.

"يبدو ذلك واقعياً جداً، لكن إرسال نمط التداخل ذلك بطريقة تمكنه من عرض مصوّر تجسيمي، هو حقيقةً صعب التنفيذ،" يقول جود. "لم يجر تنفيذه إلا في المخابر." كما أنّ المصوّرات التجسيمية الحقيقية لا تظهر ملونة كلياً، يقول جود، لأنها مولدة من حزمة ليزر صافية جداً بلون محدد، لذلك تكون عادةً خضراء أو حمراء.

إن ما يسمى تجارياً بنظام إسقاط تصوير تجسيمي هو في الحقيقة ليس إلا "تسويقاً جيداً، ويبدو ممتازاً"، يقول جود. يؤكد "وادواني" أنّ نظام إسقاط التصوير التجسيمي العالي الدقة الخاص بـ AV Concepts يعتمد على عرض شبح الفلفل، شأنه بذلك شأن نظام Musion. "لقد حدثنا بتقانة القرن الحادي والعشرين لإنشاء عرض ثلاثي الأبعاد بالحجم الحقيقي يتحرك ويتفاعل ضمن إعدادات منصة حية،" يضيف وادواني. "يتطلب إنشاء عروض تصوير تجسيمي رقائقي خاصة مشدودة جيداً، وتأثيرات ضوئية ... ومخدم إسقاط للعرض التصويري المتغير بسهولة لإنشاء خبرات مرئية ديناميكية."

تؤدي البرمجيات والعنديات دوراً أساسياً في إنتاج العروض وتصميمها وإنشائها وتنفيذها، كما يقول. يُضاف إلى قدرة المعالجة المرئية وحدات معالجة مركزية قوية، ومعايير البرمجة التفرعية متزايدة الشهرة من قبيل OpenCL، وهذا يعطي شركة AV Concepts القدرة على إنشاء عروض ضخمة جداً غامرة و"أسرة على نحو لا يُصدّق لم تكن ممكنة من قبل." ومع ذلك، يضيف وادواني، يمكن ظهور تعقيدات عند محاولة إنشاء هذا النوع من الوسائط التي تحتاج إلى معالجة قوية وضمان موثوقيتها، وعملها في الزمن الحقيقي، وجودتها العالية.

"في بعض الأحيان، نعرض عشرات الملايين من البيكسلات من المحتوى الفيديوي غير المضغوط لكل وحدة عرض ومعالجة. لا يتطلب ذلك متطلبات معالجة مرئية ضخمة فحسب، بل أيضاً تبعات الحجم الضخمة من المعطيات، مما يدفعنا لاستعمال مجموعات من الأقراص الصلبة ذات الحجم الكبير وفق تقانة RAID،" يقول وادواني. "يُنشئ تشبيك هذه الأنظمة معاً بنية مخدم قوي يمكن مكاملته بموثوقية ضمن بيئات شديدة التحدي."

التصوير التجسيمي باعتباره فناً

درس إيكو ناكامورا Ikuo Nakamura، مؤسس استديو Hololab Studio في مدينة نيويورك، التصوير التجسيمي خلال دراسته للفيزياء في جامعة طوكيو للعلوم Tokyo University of Science. يستعمل ناكامورا نبضة ليزر ياقوتية لتصوير

اللوحات أو الأغراض الحية، وقد أنشأ معدات تصوير تجسيمي (<http://www.hololab.com/holographyart.php>) حصل بموجبها على جوائز عديدة.

يعمل ناكامورا مع الأفلام الرقمية والثلاثية الأبعاد، ويقول أن الحواسيب تؤدي دوراً متمماً في التصوير التجسيمي، لأنها تحتفظ بذواكر. ويقول "الحاسوب بالنسبة لي أداة توسيع تُستعمل "كجهاز تفاعلي". ويقول أيضاً "أريد أن أوضح، الإسقاط على شاشات شفافة ليس تصويراً تجسيمياً. التصوير التجسيمي له حجم ثلاثي الأبعاد حقيقي".

النظر إلى الأمام

بمصطلحات إنشاء مصوّرات تجسيمية مولدة حاسوبياً، من الممكن الآن توليد ما يسميه جود "تمط تداخلي تركيبية"، لكن دقته منخفضة جداً. "سوف يمكن فعلياً توليد إسقاطات معقدة تماماً، لكن مقدار الحوسبة [القدرة المطلوبة] ستكون حتماً كبيرة"، كما يقول.

يقول وادواني مع تقدم التقنية فإن الإمكانيات لا نهائية. ويقول إن شركة AV Concepts طوّرت حديثاً لوحات عرض تصوير تجسيمي جديدة قابلة للتوسعة وقادرة على إنشاء عروض تصوير تجسيمي على مقياس أصغر بكثير بسرعة وبسهولة وبسعر مناسب - وهو شيء تطلب تاريخياً وتكنولوجياً وبرمجيات وهندسة عالية ومخصصة.

الخطط المستقبلية لـ Eyeliner هي "جعل الصورة أكثر واقعية، وذلك بالجمع بين فيديو 4k وإسقاط صور أكثر سطوعاً باستعمال تقنية LED من الجيل الأخير".

في هذه الأثناء، يقول أوكونيل تعمل شركة Musion على إعادة إحياء الفنان الميت التالي باستعمال تقنية Eyeliner. "كل ما أستطيع قوله هو أنّ الفنان ذكّر، وهو مازال نجماً كبيراً اليوم وربما يقدم حفلاته الأولى في الولايات المتحدة في حدود أيار أو حزيران هذه السنة (2014)".

للقراءة عن الموضوع

- Tsang, P.W. M. Liu, J. P., Cheung, K. W. K. Poon, T. C., Modern method for fast generation of digital holograms. 3D Research Center and Springer, 2010.
- Tsang, Peter, Cheung, W.K., Poon, T.C., Zhou, Z, Holographic video at 40 frames per second for 4 million object points. Department of Electronic Engineering, City University of Hong Kong, Hong Kong; Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA.; and Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Science, Shanghai, Optical Society of America, 2011.
- Tsang, P.W. M. and Poon, T. C., Review on theory and applications of wavefront recording plane framework in generation and processing of digital holograms. Department of Electronic Engineering, City University of Hong Kong, Hong Kong; Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA. Chinese Optics Letters, 2013.
- Minos K. Balvan, Object image correction using an X-ray dynamical diffraction Fraunhofer hologram, Journal of Synchrotron Radiation, Vol. 21, Part 2. Epub March, 2014.

- Bergstrom, P., Khodadad , D., Hollstig, E., Sjodahl, M., Dual-wavelength digital holography: Single-shot shape evaluation using speckle displacements and regularization. U.S. National Library of Medicine, National Institute of Health. Applied Optics. January, 2014.
- Smalley, D. E., Smithwick, Q.Y.J., Bave, V.M., Barbas, J., Jolly, S., Anisotropic Leaky-mode modulator for holographic video displays. Nature, 498, 313-317. June, 2013.

المؤلف

إيستر شين Esther Shein هو كاتب مستقل في التقانة والأعمال مقيم في منطقة بوسطن Boston.

قراءة الأدمغة

READING BRAINS*

Erica Klarreich
ترجمة: د. ندى غنيم
مراجعة: د. طلال الشهابي

جرى اتخاذ الخطوات الأولى نحو تمكين الحاسوب من إدراك أفكار أهدنا.

تقليدياً، كانت قراءة العقل مجال الصوفيين (mystics) وكُتِّب الخيال العلمي، غير أنها بدأت تصبح على نحو متزايد أحد مجالات العلم الجاد. تُبينُ دراسةً حديثةً أجريت في مخبر مارسيل فان غريفين Marcel van Gerven في مدرسة نيميغن للإدارة في جامعة رادبود (Radboud University) Nijmegen School of Management في هولندا أن بالإمكان اكتشاف ما ينظر إليه الناس عن طريق مسح أدمغتهم. عندما نظرَ متطوعون إلى حروفٍ مكتوبةٍ بخط اليد، تمكَّنَ نموذجٌ حاسوبيٌّ من توليد صورٍ مشوشةٍ fuzzy للحروف التي كانوا يرونها، بالاعتماد على نشاط دماغ المتطوعين فقط. يقول جاك غالانت Jack Gallant، وهو عالم أعصاب في جامعة كاليفورنيا في بيركلي، أن هذا العمل الجديد - المبني على نموذج رياضي سابقٍ وضعه برتراند ثيريون Bertrand Thirion من "معهد البحث في علوم الحاسوب والتحكم" في جيف سور إيفيت Gif-sur-Yvette في فرنسا- يؤسسُ خوارزميةً فكِّ ترميزٍ بسيطةٍ وأنيقة. ووفقاً لجاك، يمكن بالنهاية استعمال خوارزميات فك الترميز هذه لإنشاء واجهات أكثر تطوراً بين الدماغ والآلة لتمكين الأشخاص ذوي الإعاقات العصبية من التعامل مع الحواسيب والآلات عن طريق أفكارهم. مع تطور التقانة، يتوقع غالانت أن يصبح ممكناً بالنهاية استعمال هذا النوع من الخوارزميات لفك ترميز الأفكار والتصورات، وربما الأحلام، حيث يقول "أنا أعتقد أنه سيكون هناك، بالنهاية، شيء يشبه مسدساً رادارياً يمكنك بمجرد تصويبه على رأس شخصٍ أن تُفكَّ ترميزَ حالته العقلية". يضيف غالانت "لدينا الإطار الرياضي التي نحتاج إليه عموماً، ومن ثم فإن القيد الرئيسي الوحيد هو مدى الجودة التي نستطيع فيها قياس نشاط الدماغ".

نموذج بسيط

في الدراسة الجديدة، التي من المقرر أن تظهر في العدد القادم من مجلة Neuroimage، نظرَ المتطوعون إلى نسخٍ مكتوبةٍ بخط اليد من الحروف B، R، A، I، N، S، وفي أثناء ذلك كانت آلة التصوير بالرنين المغناطيسي الوظيفي (fMRI) functional magnetic resonance imaging تقيس استجابات قشرتهم البصرية الأولية primary visual

* نُشر هذا البحث في مجلة Communications of the ACM، المجلد 57، العدد 3، آذار (مارس) 2014، الصفحات 12 - 14.

cortex (V1) - وهي منطقة في الدماغ تقوم بالمعالجة الأولية المنخفضة المستوى للمعلومات البصرية. استعمل فريق البحث لاحقاً معطيات صور fMRI هذه لتدريب نموذج بايز Bayesian حاسوبي على قراءة عقول المتطوعين عندما عُرضت عليهم عيناتٌ جديدةٌ من الأحرف الستة.

يقول توماس ناسلاريس Thomas Naselaris، وهو عالم أعصابٍ في كلية الطب بجامعة ساوث كارولينا في تشارلستون، "إنها دراسة رائعة جداً".

وفقاً لقانون بايز "Bayes' Law"، تتطلبُ إعادةُ بناءِ الصورةِ المكتوبةِ بخط اليد التي من المحتمل أن تكون قد وُلدت نمطاً محدداً من النشاط الدماغى، معرفةً شيئين عن كل صورة مرشحة: "النموذج الأمامى forward model"، وهو احتمال



مريض يرتدي قبعة مرصعة بمسارٍ (أقطاب كهربائية) خلال استعراض المعهد الاتحادي السويسري للتقانة في لوزان لواجهة غير دخيلة بين الدماغ والآلة في كانون الثاني (يناير)

أن تولد الصورة المرشحة هذا النمط الدماغى المحدد؛ و"القبليّ prior"، وهو احتمال أن تَرِدَ هذه الصورة المحددة في مجموعة حروفٍ مكتوبةٍ بخط اليد. ومن ثمّ، فإن أية صورةٍ مرشحةٍ، تجعل جداءً هذين الاحتمالين أعظماً، هي على الأغلب الصورة التي رآها الشخص.

لإنشاء النموذج الأمامى، قام فريقُ البحث بعرضِ مئاتٍ من الصور المختلفة للحروف الستة المكتوبة بخط اليد على متطوعين، مع قياس نشاطهم الدماغى في الوقت ذاته، ثم استعملوا تقنيات التعلم الآلى لنمذجة نمطِ الدماغ الأرحج الذي ستولده أية صورة جديدة. لبناء القبليّ، أعد الفريقُ خوارزميات التعلم الآلى ثانياً للعمل على 700 نسخة إضافية من كل حرف، وذلك بهدف توليد نموذجٍ لتوضعات البكسلات الأرحج عندما يكتب الناس حرفاً ما باليد. استعمل كلا النموذجين

توزيعات غاوس الاحتمالية Gaussian probability distributions الخطية البسيطة، وهذا ما حوّل، وفقاً للباحث فان غريفين، فكّ ترميز الدماغ إلى عملية حسابية مباشرة.

يقول الباحث أيضاً: "لقد أظهرنا أن النماذج الحسابية البسيطة يمكن أن تكون تجسيداً جيداً"

أجرى فريقُ البحث أيضاً تجارب حدّ فيها من معرفة النموذج القبليّة عن عالم الحروف المكتوبة بخط اليد. فعلى سبيل المثال، إذا تكونت المعلومات القبليّة للنموذج من صور الحروف A، I، N، S فقط، فإنّ بإمكان النموذج أن يولد تجسيدات reconstructions مقبولة للحرف B، لكن ليس بالجودة التي نحصل عليها عند تضمّن القبلي صور الحروف الستة كلها. يقول فان غريفين أن هذه النتائج تظهر قدرة خوارزمية فك الترميز على التعميم -إعادة بناء أنواع من الحروف لم يسبق للخوارزمية "مشاهدتها".

يعدّ الدماغ البشري، بلا ريب، سيّد هذا النوع من التعميم، وتذهب هذه القدرة إلى أبعد بكثير من إعادة بناء بسيطةٍ لصورٍ غير مألوفة. يقول ناسيلاريس "يمكن للنظام البصري أن يفعل شيئاً لا يمكن لأي ريبوت robot القيام به، إذ يمكنه أن يتجول في غرفةٍ مملوءة بأشياء لم يرها من قبل ويتعرّف كل شيء فيها ويفهم معنى ذلك كله."

وفي حين لا يعالجُ مقال فان غريفين إلا موضوع إعادة بناء الصورة التي رآها شخص ما، اتخذَ باحثون آخرون خطواتٍ أولى نحو فكّ ترميز المعاني التي يُرَفِّقها الدماغُ بالمحفّزات البصرية، فعلى سبيل المثال، دمَجَ فريقُ غالانت (وفيهم ناسيلاريس، وهو باحثٌ في مرحلة ما بعد الدكتوراه postdoc في بيركلي) معطياتٍ من منطقة V1 (القشرة البصرية الأولية) ومن مناطق ذات درجة أعلى في المعالجة البصرية، وذلك لإعادة بناء كلِّ من الصورة التي رآها الشخص وتفسير الدماغ للأشياء الموجودة في الصورة. طبّقَ الفريقُ، في الآونة الأخيرة، في عملٍ غير منشور جزئياً، الشيء نفسه على الأفلام بدلاً من الصور الثابتة.

يقول ناسيلاريس "نحن نبدأ ببناء ذخيرة من النماذج التي يمكنها التنبؤ بما يحدث في المستويات العليا من هرمية الرؤية، حيث يجري تعرّف الأشياء".

يعمل باحثون آخرون على قراءة أفكار الدماغ عندما يستجيبُ لمحفّزاتٍ لفظيةٍ، فعلى سبيل المثال، طوّر مخبر توم ميتشل Tom Mitchell في جامعة كارنيجي ميلون Carnegie Mellon في مدينة بيتسبرغ عام 2010، نموذجاً يمكنه إعادة بناء الاسم الذي كان يقرأه شخص ما. حالياً، يعملُ مخبرُ فان غريفين على فكّ ترميز المفاهيم التي يفكر بها متطوعون عند استماعهم إلى قصة أطفال أثناء وجودهم داخل الماسحة fMRI.

سبر الأفكار

ركّزت معظم بحوث قراءة العقول حتى الآن على إعادة بناء المحفزات الخارجية التي تولد نمطاً معيناً من النشاط الدماغي. السؤال المنطقي هنا هو هل يمكنُ لخوارزميات فكّ ترميز الدماغ تحقيقَ القفزة لإعادة بناء الأفكار والتصورات الخاصة بشخص ما، في غياب أي محفّز معين.

تعتدُّ الإجابة على الدرجة التي يعالج فيها الدماغُ الصورَ الذهنيةَ والصورَ الحقيقية، مثلاً، بالطريقة نفسها. يقول فان غريفين "الفرضية هي أنّ التصور perception والصور imagery يُنشّطان مناطق الدماغ نفسها بطرق متماثلة"، وبضيف "هناك دلائل على أنّ ذلك هو غالباً ما يحصل، لكننا لم نصل إلى إثبات ذلك بعد".

يقول ناسيلريس، "إذا جرى استدعاء الإجراءات البصرية العليا بمجرد تفكيرك بشيء ما - مثل التخطيط ليومك - عندها ينبغي أن يكون بالإمكان تطوير مجسات حساسة للأفكار الداخلية، والقيام بما يشبه قراءة العقل عن طريق معرفة كيف تعمل VI فقط". ولكن/إنا هذه مسألة كبيرة".

حتى لو تبين أن قراءة العقل ليست ببساطة فك ترميز VI، يتوقع ناسيلريس أنه حالما يطور علماء الأعصاب نماذج أمامية (forward models) لمناطق المعالجة العالية المستوى في الدماغ، فإن نماذج فك الترميز ستوفر، بصورة شبه مؤكدة، بوابة إلى أفكار الناس. ويضيف ناسيلريس "لا أعتقد أن هناك استشرافاً مستقبلياً يتعلق بفكرة أنه في غضون مدة 5-20 سنة سنكون قادرين على إنشاء صور لما يفكر به الناس، أو على نسخ الكلمات التي يقولها الناس لأنفسهم". يقول غالانت إن المسألة التي قد تكون أكثر صعوبة هي نبش الذكريات القديمة للشخص أو روابط اللاوعي لديه، ويوضح فيقول "إذا طلبت منك اسم معلمك في الصف الأول، فقد تتمكن من تذكره، ولكننا لا نفهم كيف يجري تخزينه أو تمثيله"، ثم يضيف "سيكون باستطاعتنا، في المستقبل القريب، فك ترميز مواضيع الساعة التي تفكر بها الآن".

يقول ناسيلريس "من المرجح أن يكون فك ترميز الأحلام تحدياً رئيسياً آخر، فهناك الكثير من الأمور التي لا نفهمها عن النوم"، ويضيف "إن فك ترميز الأحلام سيحصل باعتقادي في المستقبل البعيد جداً". يقول غالانت "إن جزءاً من المشكلة يتلخص في أنه ليس هناك حقيقة واضحة في حالة الأحلام"، عندما يتعلق الأمر ببناء نموذج للأفكار أو الرؤى في حالة اليقظة، فمن الممكن دوماً أن تسأل الشخص عما يفكر به، أو أن تتحكم بالمحفزات التي يتلقاها دماغ الشخص مباشرة، ولكن ليس بالإمكان التحقق من حقيقة الأحلام.

لذلك، فإن الخيار الرئيسي المتاح للباحثين هو بناء نماذج لإعادة بناء الأفلام، ثم معالجة الحلم كما لو كان فيلماً يُعرض في دماغ الشخص. يقول غالانت "هذا ليس نموذجاً صحيحاً، ولكننا نستعمله على أية حال، ثم يضيف "لن يكون النموذج دقيقاً جداً، ولكن لما كانت الدقة الحالية معدومة، فإن الحصول على دقة رديئة أفضل من لا شيء".

في أيار من عام 2013، نشر فريق في مخابر علم الأعصاب الحسابي Computational Neuroscience Laboratories (ضمن معهد البحث المتقدم في الاتصالات من بعد Advanced Telecommunications Research Institute International (ATR)) في كيوتو، اليابان، برئاسة يوكياسو كاميتاني Yukiyasu Kamitani، دراسة استعملوا فيها معطيات المسح بالرنين المغناطيسي الوظيفي fMRI لإعادة بناء أصناف الأشياء المرئية التي شهدتها الأشخاص خلال الأحلام في مرحلة النعاس hypnagogic dreams، وهي الأحلام التي تحدث عندما يغلب النوم شخصاً ما". يقول غالانت "إنها ليست أحلاماً حقيقية، لكن ذلك يعتبر إثباتاً لصحة فكرة أنه سيكون بالإمكان فك ترميز الأحلام".

حماية الخصوصية

إن المستقبل البائس الذي يتصوره غالانت، والذي نستطيع فيه قراءة أفكار الآخرين الخاصة باستعمال ما يشبه مسدس الرادار، لن يحدث في وقت قريب. إن أفضل أداة متاحة حالياً للباحثين، أي الماسحة fMRI، هي في أحسن الحالات أداة غير دقيقة، إذ يمكنها فقط الكشف عن تدفق الدم في الدماغ - وهو ما يسميه غالانت "أصداء النشاط العصبي" - بدلاً من قياس الاستجابات العصبية مباشرة. إن إعادة البناء الناتج هي ظلال غامضة للمحفزات الأصلية.

من جهة أخرى، تُعد قراءة العقل باعتماد التصوير بالرنين المغناطيسي الوظيفي عملية باهظة الثمن، ومنخفضة الدقة، وغير قابلة للحمل، كما يمكن إفسالها بسهولة، حيث يقول ناسيلريس "إذا لم أكن أريد أن يقرأ أحد ما يدور في عقلي،

فيمكنني الحؤول دون ذلك"، ويضيف "من السهل أن تولّد إشاراتٍ ضجيجيةٍ في صور الرنين المغناطيسي؛ إذ يكفي أن تُحرّك رأسك، أن تغمز، أن تفكّر في أمورٍ أخرى، أو أن تذهب إلى النوم. كذلك، فإنّ هذه القيود تجعلُ التصويرَ بالرنين المغناطيسي الوظيفي أداةً غير فعّالةٍ لمعظم أنواع الواجهات بين الآلة والدماغ. يقول غالانت "يمكن تصور أنه قد يجري بالنهاية استعمال التصوير بالرنين المغناطيسي الوظيفي لتمكين الأطباء من قراءة أفكار المرضى غير القادرين على الكلام، إلا أنّ معظم تطبيقات الواجهات بين الآلة والدماغ تتطلبُ تقانةً أكثر قابليةً للحمل من التصوير بالرنين المغناطيسي الوظيفي. مع ذلك، بالنظر إلى الوتيرة الاستثنائية التي تتطور فيها التكنولوجيا، يتوقّع غالانت أن تحلّ أدواتٌ أشدّ فاعليةً محلّ الرنين المغناطيسي الوظيفي بعد مدةٍ ليست بالطويلة. يقول غالانت "عندما يحدث ذلك، ستوظفُ خوارزمياتُ فكِّ ترميز الدماغ التي وضعها ثيريون، وفان غريفين، وغيرهما، مباشرةً في التقانة الحديثة"، ويضيف "يبقى الإطار الرياضي هو نفسه إلى حدٍ بعيد، بصرف النظر عن كيفية قياسنا لنشاط الدماغ" يشعُرُ غالانت بالقلق من فكرة وجود تقانةٍ محمولةٍ لقراءة الأفكار برغم فائدتها المحتملة للمرضى الذين يحتاجون إلى واجهاتٍ بين الآلة والدماغ، حيث يقول "إنه أمر مخيف جداً، لكنه سيحدث"، ويضيف "نحن بحاجة لوضع إرشادات تتعلق بالخصوصية الآن، قبل أن تصبح التقانة على الخط.

قراءة للاستزادة

- *Horikawa, T., Tamaki, M., Miyawaki, Y., Kamitani, Y.* Neural Decoding of Visual Imagery During Sleep, *Science* Vol. 340 No., 6132, 639-642, 3 May 2013.
- *Kay, K. N., Naselaris, T., Prenger, R. J., Gallant, J. L.* Identifying Natural Images from Human Brain Activity, *Nature* 452, 352-355, March 20, 2008. <http://www.ncbi.nlm.nih.gov/pubmed/18322462>.
- *Mitchell, T., Shinkareva, S., Carlson, A., Chang, K-M., Malave, V., Mason, R., Just, M.* Predicting Human Brain Activity Associated with the Meanings of Nouns, *Science* Vol. 320 No. 5880, 1191-1195, 30 May 2008. <http://www.sciencemag.org/content/320/5880/1191>.
- *Nishimoto, S., Vu, A. T., Naselaris, T., Benjamini, Y., Yu, B., Gallant, J. L.* Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies, *Current Biology* Vol. 21 Issue 19, 1641-1646, 11 October 2011. <http://www.sciencedirect.com/science/article/pii/S0960982211009377>.
- *Schoenmakers, S., Barth, M., Heskes, T. van Gerven, M.* Linear Reconstruction of Perceived Images from Human Brain Activity, *Neuroimage* 83, 951-61, December 2013. <http://www.ncbi.nlm.nih.gov/pubmed/23886984>.
- *Thirion, B., Duchesnay, E., Hubbard, E., Dubois, J., Poline, J. B., Lebihan, D., Dehaene, S.* Inverse Retinotopy: Inferring the Visual Content of Images from Brain Activation Patterns, *Neuroimage* 33, 1104-1116, December 2006. <http://www.ncbi.nlm.nih.gov/pubmed/17029988>.

استعمال معطيات المريض لإجراء معالجات سرطان مُخصّصة

USING PATIENT DATA FOR PERSONALIZED CANCER TREATMENTS*

Chris Edwards

ترجمة: د. غيداء ريداوي
مراجعة: د. محمد عباسي

في نهاية المطاف، ستساعد قواعد معطيات المعلومات المتعلقة بالمرضى على تحسين النتائج الصحية وعلى دعم تطوير علاجات جديدة.

تمضي منظمات الصحة في أوروبا والولايات المتحدة الأمريكية قدماً في تجميع قواعد معطيات ضخمة لمعلومات المرضى وفق خططٍ معدّة، وتتوقّع ألا يقتصر أثر هذا التجميع على تحسين المعالجة وفرص النجاة لمن يعانون من السرطان وحسب، بل يتعداه إلى تطوير علاجات جديدة. سيفوق عدد المشاركين في هذه الخطوات عدد من يشاركون حالياً في تجارب سريرية واسعة النطاق، وستتمكّن المراكز الطبية في الدول المتطورة من توسيع انتشارها عالمياً. تهدف إحدى المنظمات، وهي الجمعية الأمريكية لعلم الأورام السريري American Society of Clinical Oncology (ASCO) إلى أن يكون الإصدار الإنتاجي production version من نظام التعلّم الصحي الخاص بها والمسمى CancerLinQ جاهزاً بحلول العام 2015. بُني هذا النظام على خبرة مستقاة من نموذج أولي يتضمن 170000 سجلّ كُشف عنه النقاب في 2013. في إحدى مناسبات البيت الأبيض White House في تشرين الثاني من العام 2013، عدّ توماس كيليل Thomas kalil - نائب مدير التقانة والإبداع في مكتب البيت الأبيض لسياسة العلوم والتقانة - النظام CancerLinQ أحد أنظمة الرعاية الصحية المتعددة التي من شأنها تجميع كميات ضخمة من المعطيات بهدف تحسين المعالجة.

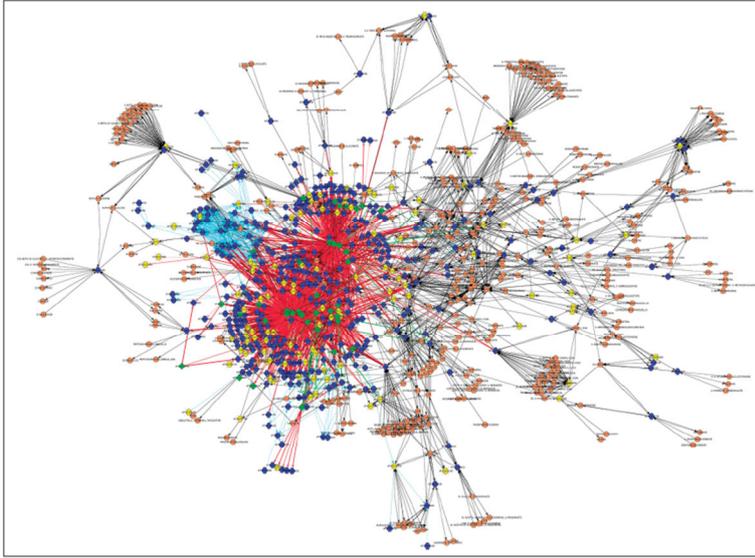
يقول كيليل "توجد فرصٌ ضخمة لتحسين النتائج الصحية وخفض التكاليف في آن واحدٍ"، مدّعياً أن النظام CancerLinQ سيسمح بأن تساعد كل تجربة من تجارب المرضى على إعلام المُتولّين أمر مرضى السرطان في المستقبل. يتابع كيليل القول "يساهم حالياً 3% فقط من المرضى في التجارب السريرية. وتلتزم الجمعية الأمريكية لعلم

* تُشر هذا البحث في مجلة Communications of the ACM، المجلد 57، العدد 4، نيسان (أبريل) 2014، الصفحات 13 - 15.

الأورام السريري ASCO بالكشف عن كيفية استعمال معطياتهم، على حين تحمي خصوصية المرضى الـ 97% الباقين، وذلك لتجعل نظام العناية الصحية أكثر بكثير من نظام تعلّم، ولتحسّن جودة الرعاية وتسرع ابتكار علاجات جديدة." يقول الدكتور بيتر كيمبل رئيس علم وراثّة السرطان والجينوم في معهد ويلكوم ترست سانجر Wellcome Trust Sanger Institute في منتدى 2013 حول سياسة الأدوية السرطانية المنعقد في أمستردام-هولندا في الخريف الماضي، إن تجميع المعطيات ضمن قواعد معطيات ضخمة متاحة على نحو مباشر قد يساعد على ظهور الطب المُخصّص، فيمنح فرصة لعلاج الأمراض الخطيرة علاجاً أشدّ فعالية. "إننا نقف على عتبة عصرٍ يمكننا فيه أن نعرف خصائص كل مريض والسبب الذي أدى إلى إصابته بالسرطان."

تُخذ السرطان هدفاً رئيساً، وذلك جزئياً بسبب انتشاره. فحسب جمعية السرطان الأمريكية، لا يزال السرطان يحتل المرتبة الثانية للوفيات في الولايات المتحدة، وبمعدل واحدة من كل أربع وفيات. يشبه الوضع في أوروبا هذا الوضع، إذ يقول تونيو بورغ مفوض الصحة في الاتحاد الأوروبي "نتوقع في أوروبا أن يصاب رجل من كل ثلاثة رجال وامرأة من كل أربع نساء بالسرطان قبل بلوغ سن الخامسة والسبعين. فالسرطان لا يصيب الآخرين فقط، بل قد يصيب أيّاً كان." تعدّ طبيعة السرطان سبباً آخر لاستعمال قواعد معطيات يتناسب حجمها مع عدد السكان لمقارنة معلومات المرضى ومعالجتها. يقول كيمبل إن للسرطان تنوعاً ضخماً ناتجاً عن الطرائق العديدة المختلفة التي يتحول فيها الحمض النووي DNA لخلايا الورم لدى المرضى المختلفين. ففي سرطان الثدي مثلاً، يكشف تحليل الأورام أن أكثر أنواعه شيوعاً، وهي الناتجة عن جينات متحوّلة (طافرة) أو محذوفة، موجودة عند 10% فقط من المرضى المصابين. ويشرح قائلاً "ثمة قائمة طويلة من التحولات الأخرى التي يصيب كل منها نسبة ضئيلة من المرضى."

في أمراض كمرض السرطان، تتسبب التغيرات الجينية بحلقات من ردود فعل حيوية وعمليات أخرى لكسر هذه الحلقات بطرائق غير متوقعة. وجد الباحثون أن الخلايا قد تتبدل بطرائق مختلفة حتى ضمن الورم نفسه، ومن ثم فإن



في 2011، بنى باحثون في المركز الطبي لجامعة كولومبيا هذا النموذج لشبكة تنظيم الجينات في خلايا الثدييات التي استخدموها في دراسات تتعلق بالتغير الجيني للسرطان.

العلاج الذي ينجح لدى مجموعة كبيرة من الأشخاص قد يكون غير مُجدّب البتة لدى أحد المرضى الجالسين في عيادة الطبيب.

يقول كيمبل إن حلم الطب المُخصّص أو الطبقي، هو أن توصف علاجات مُخصّصة لحالة المريض "لجعل السرطان يتلاشى نهائياً". تكمن المشكلة في الحصول على ما يكفي من المعطيات لتحديد مدى نجاح أو إخفاق العلاجات المختلفة ضمن الشروط المختلفة.

يشرح كيمبل قائلاً "أود أن أقول إن التجارب السريرية المتداخلة لكشف

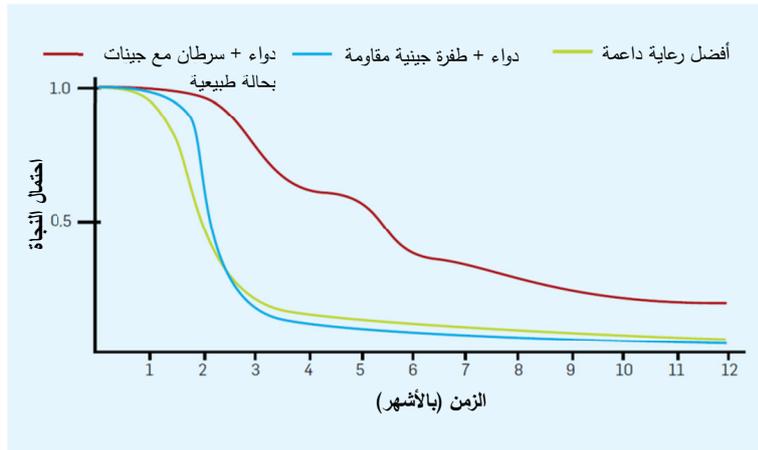
التفاعل بين الجينات والأدوية قد ضَعُفَت. فإن وجد تحول (أو طفرة) محدد في مريض واحد فقط ضمن 700 مريض، فإنك تحتاج إلى فحص هذا الكَم لإيجاد مشارك واحد في تجربة سريرية لتجريب دواء يستهدف هذا التحول". وبضيف، "لا يوجد لدينا العدد الكافي من المرضى لدراسته. نحتاج إلى عدة آلاف من المرضى، ربما عشرات الآلاف. ولكل نوع معطى من أنواع الأورام نحتاج إلى قاعدة معطيات من 10000 إلى 20000 مريض، وبوجود 50 إلى 100 نوع معروف من الأورام نحتاج إلى النفاذ إلى سجلات مليون مريض على الأقل".

تقول الدكتورة ساندرا سوين الرئيسة السابقة للجمعية الأمريكية لعلم الأورام السريري ASCO، والمدير الطبي لمعهد واشنطن للسرطان في مركز واشنطن للاستشفاء "مدستار" MedStar، إنه توجد مجموعات كبيرة من المرضى غير ممثلة جيداً في بنية التجارب السريرية الحالية. فالكبار مثلاً، "يعانون من أمراض مرتبطة بتقدمهم في العمر؛ وهذا مما يغير من كيفية قيامنا بتقييم المريض الذي يقف أمامنا".

نقطة هامة أخرى هي كيف يتمكن الأطباء الممارسون من التدقيق في المعطيات المتاحة نظرياً، ويتعدَّر النفاذ إليها عملياً.

لتحسين إمكان نفاذ الأطباء الممارسين إلى المعطيات الطبية المناسبة يستعمل مركز السرطان إم دي أندرسون في جامعة تكساس حاسوباً فائقاً من نوع أي بي إم تعمل عليه برمجيات للذكاء الصناعي كتبت في الأصل لتمكين هذا الجهاز الضخم من الدخول في المنافسة في عرض لعبة "جابردي! Jeopardy!". يستعمل الحاسوب الفائق واتسون أي بي إم Watson IBM تقنيات إحصائية مستعملة في معالجة اللغات الطبيعية لكشف أجزاء المعلومات التي تتعلق بمسألة ما. يجري استعمال برمجية "مرشد خبير الأورام" (OEA) Oncology Expert Adviser في تحليل معطيات المريض، كما تستعمل معطيات بحثية من تجارب سرطانية أُجريت في المركز.

يقول الدكتور هاكوب قنطارجيان رئيس قسم اللوكيميا وأستاذ اللوكيميا في مؤسسة إم دي أندرسون إن النظام البرمجي المعتمد على الحاسوب واتسون في المركز "سيكون أداة تحوُّل. فعندما نرى اليوم المرضى المصابين بالسرطان،



منحني كبلان-مير لفرص النجاة مع الزمن، ويستخدم عموماً لقياس فعالية العلاج عند المرضى، وهو يُظهر كيف يقوم السرطان المتحوُّل من شكل يمكن مصادفته بكثرة بتطوير ممانعة عالية للمعالجة الدوائية المعيارية.

فإننا نعتد على ذاكرتنا ومعرفتنا المحدودة لتعيين الخطوة التالية الفضلى. قد يصبح

"مرشد خبير الأورام" OEA، بسبب عرضه للمعلومات من مصادر متعددة، نقلةً نوعيةً في الطريقة التي تُعالج بها المرضى ونقدّم بها الرعاية والمعرفة".

استعملت الجمعية الأمريكية لعلم

الأورام السريري ASCO في نموذجها

الأولي CancerLinQ مقارنة معتمدة

على تشكيلة من البرمجيات المفتوحة

المصدر وبرمجية التنقيب في المعطيات،

غاليليو كوسموس Galileo Cosmos،

التي تقدمها تحليلات غاليليو Galileo

Analytics، إذ نفذت إلى المعطيات المعالجة قبلاً وحللتها. تستعمل برمجياتُ المعالجة القبليةُ توابع إحصائية وشبكة عصبونات صُنعية للتعلّم، ولهيكلة المعطيات وتنظيم حقولها في السجلات الأصلية، وفق المصاغة التي يحتاجها النظام.

يقول الدكتور كليفورد هوديس رئيس الجمعية الأمريكية لعلم الأورام السريري ASCO ورئيس خدمة علاج سرطان الثدي في مركز سرطان ميموريال سلون-كيتيرينغ Memorial Sloan Kettering Cancer Center في نيويورك، يسمح النظام CancerLinQ للأطباء برؤية نتائج المداخلات على مرضى آخرين، يتسمون بِسِمَاتٍ (لاحقة) معينة، لمساعدتهم على تحديد مسار العلاج الأكثر ملاءمة. ويؤكد أن "الطبيب ليس مجرد ربوط ينفذ ما يطلبه منه الحاسوب".

يقول قنطاريجان عن "مرشد خبير الأورام" OEA "ستسمح أداة تجميع وتحليل المعطيات هذه باكتشافات لا يمكننا إجراؤها حالياً. لنفرض أن مجموعة من المرضى تتناول دواء ليس مخصصاً للسرطان، وليكن مثلاً دواء للقلب، وأثر هذا الدواء في مدى تحسُّس الورم لدى هؤلاء المرضى لمفعول المعالجات السرطانية. سنكتشف سريعاً، باستعمال هذه الأدوات التحليلية، أن هذا الدواء يساعد المرضى، ومن ثمّ يمكننا أن نطبقه في برامج علاجية مستقبلية".

يحدّر الدكتور إميل فوست رئيس مركز العلاج السرطاني الشخصي في معهد الأبحاث الطبية يو إم سي أوترخت UMC Utrecht في هولندا، الذي لديه نظام خاص لتحليل المعطيات، من وضع ثقة كبيرة في النتائج المستخلصة من قواعد المعطيات الكبيرة الحجم، لوصف علاجات مخصصة "دون تصريح بالاستعمال off-label"، قبل إجراء المزيد من الأبحاث.

يقول فوست "يقول بعض الأشخاص إنه يمكن معالجة 75 % من المرضى بدواء ما نظراً للمعلومات الجينية التي نمتلكها حالياً. أنا أنزع إلى الاختلاف في الرأي، لأننا بصراحة لا نعلم، يمكننا فقط أن نفترض ذلك. قد ينجح الأمر، لكنني أعارض بشدة معالجة المرضى "دون تصريح باستعمال الدواء لهذا المرض" خارج الدراسات السريرية لأننا سنتسبب بحالة فوضى".

فشلت محاولة سابقة لاستعمال المعطيات الجينية لانتقاء العلاج. بدأت التجارب السريرية في 2007 بعد أن نشر باحثو جامعة ديوك Duke نتائج تقترح اختبارات تستعمل صفيقة (مجموعة) من المسابر الكيميائية للحمض النووي الريبي يمكنها المساعدة على انتقاء العلاج القابل للتطبيق. ثم وجدت دراسة لاحقة أنه لا يمكن إعادة إنتاج تلك النتائج، وقامت مجموعة من 30 متخصصاً في المعلوماتية الحيوية والإحصاء بحضّ المعهد الوطني للسرطان (NCI) على تعليق هذه التجارب، وهذا ما فعله.

إن لم تُستعمل المعطيات التي جرى التقيب عنها استعمالاً مباشراً لابتكار معالجات مخصصة لم يصرح باستعمالها في عيادات الأطباء، فمن المرجح أن تستعمل في إعطاء معلومات للبحوث الطبية الأساسية، خصوصاً في المجال الناشئ المتعلق ببيولوجيا الأنظمة الذي يستعمل النمذجة الحاسوبية لتوقع السلوك البيولوجي.

تقول أندريا كاليانو رئيسة قسم بيولوجيا الأنظمة في جامعة كولومبيا "بدلاً من استعمال الاقترانات الإحصائية، بدأنا بالتوجه نحو نمط من العلوم مقود بالنماذج حيث يمكننا صنع نماذج تنظيمية للخلايا".

تتظر بيولوجيا الأنظمة إلى الخلايا البيولوجية كأحد أشكال شبكات الإنذار، وهي نظام ديناميكي تتفاعل داخله البروتينات والجينات في حلقات ردود فعل معقدة. يعتقد عدد من علماء السرطان أن العلاجات المركبة قد تقدم مفتاح التعاطي مع السرطانات التي يصعب حالياً علاجها مباشرةً.

يمكن للشبكة المعقدة داخل الخلية أن تقدم طريقة أشدّ فعالية لمهاجمة السرطانات. يمكن لدواء منفرد أن يستهدف جزءاً من الشبكة فقط، وربما همد بروتيناً واحداً، غير أن هذا الأمر لا يكون فعالاً إلا إذا جرى تعطيل الخلية المسرطنة في هذه الخطوة. في العديد من الحالات، تؤدي التغييرات التي تصنع في الشبكة، نتيجة إيقاف أحد البروتينات عن عمله، إلى كشف ممرات إنذار جديدة تساعد الخلية على البقاء حية. تساعد النمذجة الحاسوبية في تحديد أكثر الأهداف المنتجة وتشكيلاتها لزيادة احتمال نجاح العلاج.

ويقول كيمبل "لا يقتصر الأمر على وجود العديد من الجينات ذات العلاقة؛ إذ إن العديد منها غير قابل للتفعيل بسهولة. ويعدّ العديد منها أساسياً: أي أنك لا تستطيع تخديرها دون سُميّة. فمثلاً، تتحكم الجينات الكابتة للورم في سرطان الكلىة تحكماً شبه تام تقريباً، ومن المعروف أنه يصعب تخدير هذه الجينات.

يضيف كيمبل "توجد مسألة أخرى هي أن العديد من هذه الجينات يتفاعل بعضها مع بعض. يمكنك أن ترى أشكالاً نمطية من الجينات التي يتشارك بعضها مع بعض في التحوّل. فعندما يتحوّل أزواج من الجينات تحوّلوا واضحاً، فهي غالباً تتفاعل بقوة؛ من الواضح أن هذا الأمر هام جداً، ويحتمل أن يظهر في معظم السرطانات. سيكون لهذه الطفرات الثانوية تأثيراً قوياً في علاج المريض."

مع أن هدف أنظمة التعلم الطبية الضخمة هو الحصول على تجميع وتحليل آليين للمعطيات، فإن الحقيقة أشدّ تعقيداً. يقول الدكتور بيتر جونسون كبير الأطباء في أبحاث السرطان في المملكة المتحدة، معتمداً على خبرته في برنامج المملكة المتحدة الرائد، إنه لا يمكن ضمان الأتمتة الكاملة. ويضيف "إننا نهدف إلى أتمتة تجميع المعطيات واستخراجها، غير أن التعقيد يبلغ درجة تجعلنا نقوم بالكثير من التجميع والاستخراج يدوياً.

ومع اقتراب النظام CancerLinQ من أن يصبح مُنتجاً، تتوقع الجمعية الأمريكية لعلم الأورام السريري ASCO أن بعض المعالجة اليدوية ستكون ضرورية، على الأقل في النسخ الأولى من النظام. يقول الدكتور روبرت هاووزر وهو مدير كبير في قسم الجودة في ASCO "إن علم الأورام علم معقد، ومن ثمّ يحتاج إلى الكثير من التفسير والتدخل اليدوي في بدايته. ونعتقد أنه بمرور الزمن سنتناقص جهود التدخل اليدوي كلما تعلّم النظام.

لما كانت هذه الأنظمة تكبر لتشمل الملايين من سجلات المرضى، فإن مؤسسات مثل إم دي أندرسون MD Anderson ترى فيها وسيلة لتوسيع انتشارها عالمياً. ويشير الدكتور رونالد ديبينو رئيس إم دي أندرسون إلى تقرير حديث صادر عن المعهد الطبي قام بتصنيف الفوارق الكبيرة في رعاية السرطان على امتداد الولايات المتحدة. أما خارج الولايات المتحدة، فيقول ديبينو "لدينا صحارى طبية على الصعيد العالمي".

ويتابع قائلاً "مهمتنا هي إنهاء مرض السرطان في منطقتنا وفي أرجاء العالم. تسمح خطّتنا بجعل الرعاية في إم دي أندرسون ذات طابع ديمقراطي. يمكننا الذهاب إلى حيث يوجد المرضى، وإلى أماكن لم يقرأ الطبيب فيها أي مقال بحثي منذ 10 سنوات."

إن تحصيل المعطيات ومعالجتها هما بداية إجرائية يمكنها أن تربط بين معالجات السرطان في أرجاء العالم.

مراجع للاستزادة

- Basso, K., Margolin, A.A., Stolovitzky, G., Klein, U., Dalla-Favera, R., Califano, A. Reverse engineering of regulatory networks in human B cells, Nature Genetics 37, 382-390 (2005)

- *Gonzalez-Angulo, A.M., Hennessey, B.T.J., Mills, G.B.* Future of Personalized Medicine in Oncology: A systems Biology Approach *Journal of Clinical Oncology*, vol 28, no.16, June 1 (2010)
- *Tsimberidou, A.M., Ringborg, U., Schilsky, R.L.* Strategies to overcome clinical, regulatory, and financial challenges in the implementation of personalized medicine 2013 ASCO Educational Book <http://meetinglibrary.asco.org/EdBookTRacks/2013%20ASCO%20Annual%20Meeting> [open Access]
- *Godman, B., et al.* Personalizing health care: feasibility and future implications *BMC Medicine*, 2013, 11:179 <http://www.biomedcentral.com/1741-7015/11/179> [open Access]

المؤلف

كريس إدواردز **Chris Edwards** كاتب مقره في سري، المملكة المتحدة، يكتب في مجالات الإلكترونيات، وتقانة المعلومات، والبيولوجيا التركيبية

تعليم الحواسيب الخدع البصرية

TEACHING COMPUTERS WITH ILLUSIONS*

Esther Shein

ترجمة: أ. سعيد الأسعد

مراجعة: أ. مروان البواب

إن استقصاء الطرائق التي يمكن أن تُخدَع الرؤية البصرية عند الإنسان ينهض اليوم رِدْءًا (معيّنًا) لمطوّري الرؤية الآلية.

في وقتٍ سابقٍ من هذا العام، احتدمَ على شبكة الإنترنت جدلٌ بشأن لون ثوب! كان التساؤل على وجه التحديد: لماذا ينظر الناسُ إلى الصورة الواحدة عينيها، فيرونها على نحوٍ مختلف؟ ولو أدخلنا الحواسيبَ في الاعتبار لوجدنا أنها، خلافًا للإناسيِّ الذين يرون صورًا معيّنَةً بهيئاتٍ مختلفة، تسجّل وتعرّف الصورَ البصرية visual images على مستوى آخرٍ مختلفٍ تمامًا. ويقرّر خبراءُ البصريات أن ما يراه الإنسان تحدده الوظائفُ البيولوجية، في حين تحدّد الحواسيبُ الرؤية من قياساتٍ فيزيائية.



الصورة المرشحة من غوغل، والفائزة في مسابقة (ImageNet 2014)، تساعد الحواسيبَ على التمييز بين الأشياء. (العبارة العلوية: قُبعة ذات حافةٍ عريضة).

وإذ قد ينبئ الميدانان أحدهما عن الآخر، يرى الباحثون أن ثمة مزيدًا من العمل لا بدّ من إنجازه، بغية تلقين الحواسيب أساليبَ تحسين تعرفها الصورَ image recognition.

وتبرز أهمية تلك المساعي بإدراك مدى رغبتنا في حمل آلاتٍ من مثل الإنسالات¹ robots على رؤية العالم كما نراه نحن. يقول جيف كلون Jeff Clune (الأستاذ المساعد ومدير مختبر الذكاء الصناعي المتطوّر في جامعة وايومينغ الأمريكية): «الأمر مُجدِّ عمليًا. غرضنا أن نجعل الإنسالاتِ عونًا لنا، كأن يكون

* نُشر هذا البحث في مجلة Communications of the ACM، المجلد 58، العدد 8، آب (أغسطس) 2015، الصفحات 13 – 15.

¹ إنسالة: نحت من كلمتي إنسان + آلة، أي إنسان آلي، أو رُبوط. (المترجم)

بإمكاننا مثلاً تلقينها أن 'تدخل المطبخ وتُمسِك بمقصي وتعيده إلي'. ومن ثم يتعيّن تعليمُ الإنسالة ما هو شكل المطبخ وما هو شكل المقصّ، وكيف تصل إليهما. يجب أن تكون الإنسالة قادرةً على رؤية العالم والكانتات فيه. هناك إذن منافع كبرى نجنيها حالما نتقن الحواسيب ذلك فعلاً.»

بيد أنه مهما بلغت قدرة الآلات على تمييز الصُور وتعرّفها، فإن ثمة عاملين -كما يقول الخبراء- يسهم غيابهما في إعاقة أدائها: الخبرة experience والتطوّر evolution.

وقد باتت الحواسيبُ بارعةً في تعرّف الوجوه بخاصة، لكنها «ظلت عاجزةً عن إدراك التفاصيل الدقيقة التي نلتقطها نحن مباشرةً لدى رؤيتنا لوجهٍ ما، ونفاذنا إلى كلّ ما يتّصل به من معلومات»، حسبما يقول ديل بيرفريز Dale Purves (أستاذ البيولوجيا العصبية في جامعة ديوك الأمريكية). ومن ناحيةٍ أخرى، فإن لدى الناس «فيضاً من المعلومات المستندة إلى ما يوحي به لنا ذلك الوجه... بحيث ندرك على الفور الدلالات السلوكية لتقطيبه أو ابتسامته تظهر عليه». ويتابع قائلاً إن تحقيق ذلك في مساق الرؤية الآلية ما زال مطلباً صعباً بعيد المنال، ذلك «لأن الآلة غير قادرة حتى الآن على إدراك ما هو مهمّ لتحقيق النجاح السلوكي في العالم وما هو غير مهمّ.»

في المقابل، يلحظ بيرفريز أن نجاح الإنسان في إدراك تلك التفاصيل الدقيقة جاء حصيلةً لملايين السنين من التطوّر، إضافةً إلى الخبرة الفردية. «وقد صرّح كثيرون، وبطرائق مختلفة، أن لا شيء في البيولوجيا يحمل معنىً منطقياً إلا في ضوء التطوّر. وأعتقد شخصياً أن ذلك هو عين الصواب، غير أن الرؤية الآلية تعجز عن إدراك هذا المبدأ.»

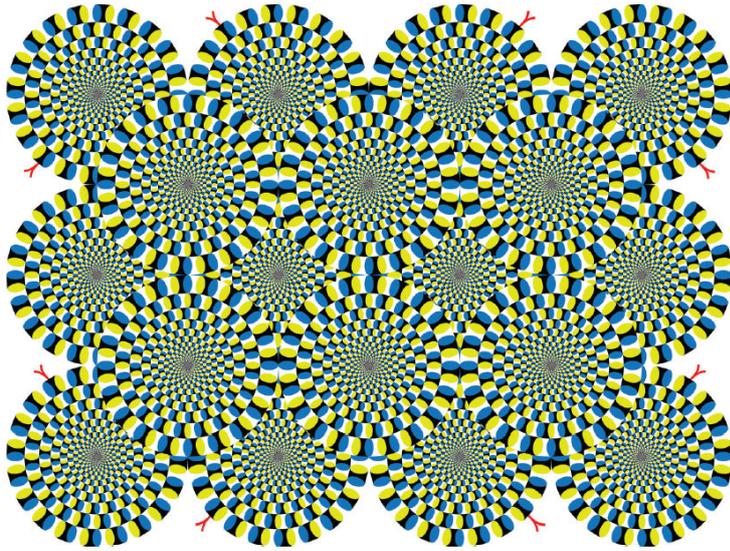
ويسعى الناس إلى جعل المنظومات الصنعية ترى العالم كما هو. «أما في حالة أدمغتنا فإن تطوّر الجملة العصبية البشرية على امتداد العصور لا يرمي بالضرورة إلى رؤية العالم كما هو - بل إلى رؤيته بطريقة تجعل استمرار وجودنا وتكاثرنا أمراً أرجح احتمالاً»، وفقاً لما تراه سوزانا مارتنيز-كونده Susana Martinez-Conde (الأستاذة الجامعية ومديرة مختبر العلوم العصبية التكاملية في المركز الطبي التابع لجامعة نيويورك الحكومية).

وتستفيد مارتنيز-كونده، التي ينحصر عملها في مجال الإدراك الحسي والانطباعات البصرية الخادعة ومنشئها العصبي، في الشرح قائلة: «إن الدماغ البشري يُجري كثيراً من التقديرات الحدسية guesstimates، فنستخلص من واقعها قدرًا محددًا من المعلومات، ثم نسدّد ونقارب في الباقي، ونعتمد إلى طرائق مختصرة للوصول إلى صورةٍ قد لا تكون تجسيداً دقيقاً لصيغها الواقعي، غير أنها تفي بالغرض.»

من الأمثلة المعروفة للانطباعات البصرية الخادعة التي يلحظها الإنسان والآلة بصورٍ مختلفة نموذج الأفاعي الدوّارة (المتحوّبة) (<http://bit.ly/1IRuVDb>). تقول مارتنيز-كونده إن الصورة في الواقع ثابتة، لكنها تبدو متحرّكة لدى معاينتها على الورق، «لأن دارات الإحساس بالحركة motion sensitivity circuits في الدماغ مركّبة، أو تعمل، بحيث تفسّر وجود تتابعٍ sequence معيّن على أنه حركة motion، حتى إن لم يكن ثمة حركة حقيقية في الصورة.»

وتضيف إن في الدماغ البشري عصبوناتٍ بصرية vision neurons تقتصر وظيفتها على اكتشاف الحركة، وهذا ما يراه غالبية الناس لدى معاينتهم الصورة. على أن للسّن دوراً فيما يراه الناس كذلك.

وتشير مارتنيز-كونده إلى أنه لما كانت خدعة الأفاعي جديدةً نسبياً، فإن الجانب الذي لم يُدرك منها بعدُ تماماً هو أن الأفراد الذين ناهزوا الأربعين من العمر أو من هم دونها هم أوفر حظاً في رؤية الحركة في الصورة، على حين ينزع الأفراد الذين بلغوا سنّ الخمسين أو جاوزوها إلى عدم القدرة على رؤيتها. ولم يُعلّم حتى الآن سببٌ للتغيّر الذي يطرأ على



خدعة الأفاعي الدوّارة (البصرية)، كما عرضها أكويشي كيتاوكا، الأستاذ في جامعة ريتسوميكان [بمدينة كيوتو اليابانية].

أداء المنظومة الحركية motor system مع تقدّم السنّ. وتقول: «من الطريف حقاً أن تعلم أنّ تراجع قدرة المنظومة البصرية الحركية motor visual system لديك مع تقدّم سنّك يجعلك أقرب إلى رؤية عالم الواقع منك إلى رؤية عالم الوهم، علماً بأن رؤيتك حركة في خدعة الأفاعي هي أمانة على سلامة حاسة بصرك.»

من ناحية أخرى، يرى بيرفيز أن الرؤية الآلية تقوم على خوارزميات تستطيع قياس مفردات في البيئة، واستعمالها في عربات تُسيّر بدون سائق، وفي أغراض أخرى. ويشار إلى أن الإنسان لا يمكنه النفاذ إلى المعلومات نفسها التي تعتمد عليها خوارزميات الآلة لغرض الرؤية.

ويرد قائلًا: «نحن البشر ... نعاني مشكلة عويصة، تتمثل في عدم قدرتنا على النفاذ إلى العالم المادي من حولنا، بسبب افتقارنا إلى طرائق لقياسه باستعمال أجهزة كالماسحات الليزرية laser scanners أو المنظومات الرادارية radar أو مقاييس الضوء الطيفي spectrophotometers أو بوسائل أخرى لإجراء قياسات لما هو موجود ماديًا. ومع ذلك، يسلم الجميع بتقوّنا على الآلة في تعرّف الوجوه وصنع القرارات باستعمال معلومات تطوّرت على مدى ملايين من السنين» على أساس من التجربة والخطأ.

على أن ذلك لم يكن ليثني الناس عن مواصلة السعي لجعل الإنسان والآلة أقرب إلى رؤية الانطباعات البصرية الخادعة نفسها. وفي سبيل تحقيق هذه الغاية يعمل كوكيشي سوغيهارا Kokichi Sugihara (أستاذ الرياضيات في جامعة مييجي Meiji في طوكيو) حاليًا على برنامج يؤمل أن يمكن الحواسيب من إدراك العمق في رسوم ثنائية البعد. ويظهر من كتابات سوغيهارا على موقع الوب الخاص بالجامعة (<http://bit.ly/1CAwz7F>) أن اهتمامه ينصب أولاً وبالذات على «تمكين الحاسوب، عن طريق معالجة معلومات الدخل، من إدراك شكل ثلاثي الأبعاد انطلاقًا من مسقط projection مرسوم بخطوط.»

وكتب سوغيهارا يقول: «غالبًا ما يخفق الحاسوب في إعادة إنشاء شكل ثلاثي الأبعاد من رسم إسقاطي projection drawing، فيعطي رسائل خطأ error messages، على حين يستطيع الإنسان تنفيذ هذه المهمة بسهولة متناهية. إننا نتخيّل جسمًا ثلاثي الأبعاد من رسم ثنائي البعد بناءً على تصوّر سابق متحصّل بفطرتنا وخبرتنا البصرية ... أما الحاسوب فلا يتأثر طبعًا بأيّ تصوّر، وإنما يختبر جميع الاحتمالات بغية إعادة إنشاء جسم ثلاثي الأبعاد، ويخلص إلى أنه قادر على ذلك.»

وثمة طرائق مختلفة يمكن استعمالها ل«خداع» الخوارزميات الحاسوبية، بحيث يكون ما تراه المنظومة وما يراه

الإنسان أكثر تقاربًا. تقول مارتينيز-كونده موضحةً ذلك: «من بين وسائل تعزيز الرؤية الصناعية artificial vision التعمق في دراسة ما تراه أدمغتنا، مع يقيننا، أولاً وأخراً، بأنها تعمل بكفاءة وبأن منظومتنا البصرية بالغة التعقيد. ومن ثم فإن تعمقنا أكثر فأكثر في إدراك هذه المنظومة من منظور عصبي من شأنه أن يحسن الرؤية الحاسوبية.» لكنها تستدرك قائلةً إن منظومتنا البصرية «ليست على الإطلاق بالغة حد الكمال؛ فلو أننا بلغنا مرحلةً كادت أن تتساوى فيها جودة الرؤية الحاسوبية والبشرية، ما كان ذلك ليعني أن العمل قد أُنجز.»

ويقول بيرفيز إن الإنسان يعمد إلى الاصطفاء الطبيعي natural selection ليُدخل في الشبكات العصبونية neural networks الموجودة في أدمغتنا كل ما يمكن تصوّره من الحالات التي تحمل دخلاً بصرياً. «وما إن تستطيع الحواسيب أن تفعل ذلك وتتطور، حتى تغدو -مبدئياً- في مستوانا نحن البشر، إلا أنها لن تضاهئنا في تحقيق القياسات البصرية؛ فهي تحقق الرؤية من طريقٍ مختلفٍ جداً. وسيكون ثمة حدٌ لن يسمح لها أبداً بالوصول إلى مستوى البشر في أداء وظائفهم.»

ومع ذلك، يمكن أن تستمر الآلات في التحسن والارتقاء. وهنا يلفت بيرفيز النظر إلى أن «السبيل إلى تحقيق آلة رصينة حقاً هو تطويرها عن طريق خبرات التجربة والخطأ trial-and-error experiences، وتصريف (ترجمة) compiling تلك الخبرات في مجموعة داراتها العصبونية الصناعية artificial neural circuitry. وليس ثمة من سبب يمنع تطبيق ذلك؛ ما عليك إلا تجميع المعلومات التي استعملناها لتطوير منظومة بصرية.» ويقدّر بيرفيز أنه في غضون عشرين سنةً يمكن أن ترتقي الرؤية الآلية لتحاكي الرؤية البشرية، أي حالما ينجح العلماء في استنباط طريقة لتطوير شبكة عصبونية صناعية artificial neural network قادرة على «البقاء والاستمرار» في بيئات «معقدة تعقيد العالم الذي نعيش فيه».

ويتفق كلون مع الرأي السابق، من أن «البشر والحواسيب ترى الأشياء على نحوٍ مختلفٍ جداً، وما زال أمامنا شوطٌ بعيدٌ من الجهد لإدراك آلية عمل هذه الشبكات.» وفي مقالةٍ بحثيةٍ تناول مسألةً مقلقةً تتمثل في أنه إذا تعرّف حاسوبٌ صورةً ساكنةً عشوائيةً، ولتكن دراجةً ناريةً مثلاً، بنسبة يقين 100%، فإنه بذلك يخلق ثغرةً أمنيةً security loophole؛ إذ كلما كان باستطاعتي أن أحمل الحاسوب على تصديق صورةٍ على أنها تمثل شيئاً ما في حين أنها في الواقع لشيءٍ آخر، توقّرت إمكاناتُ استغلال ذلك لمصالح شخصيةٍ تعود إلى شخصٍ آخر.»

فعلى سبيل المثال، ربما أنتجت شركة للفن الإباحي صوراً تبدو لمرشحات الصور image filters (العائدة لشركة Google) أشبه بأرانب، لكنها في الواقع تحتوي على إعلاناتٍ تُدسُّ في مضمونها العُري؛ أو قد تحتال مجموعةٌ إرهابيةٌ على مرشحات الذكاء الصناعي artificial intelligence filters التي تفتش عن نصٍّ متضمنٍ في صور، وذلك بجعل تلك الصور تبدو للذكاء الصناعي صوراً لأزهار. كذلك قد تعاني مظاهر الأمن الإحصائي الحيوي biometric security قصوراً وتكون عرضةً للاختراق؛ ومن أمثلة ذلك أن «يرتدي إرهابيٌّ قناعاً من غشاء بلاستيكيٍّ شفافٍ ورقيقٍ جداً طُبِعَ عليه شكلٌ مشوّش لا تراه العين البشرية، لكنه يخدع منظومةً لتعرّف الوجوه facial recognition system، بحيث ترى أمامها عميلٌ آمنٌ ذا سلطات بدلاً من أن تتعرّف إرهابياً خطراً»، حسبما يقول كلون.

ويتابع: ومن عجَبٍ ألا يكون اعتقادُ البعض بإمكان انخداع منظومةٍ ما بصورٍ معينةٍ وعدم انخداع منظومةٍ أخرى متمرسيةً على تعرّفها اعتقاداً صحيحاً دوماً؛ «ففي وسعي توليد صورٍ بشبكيةٍ ما، وعرضها على شبكيةٍ أخرى مختلفةٍ تماماً مراتٍ ومراتٍ، وفي كل مرةٍ تتخدع هذه الشبكة الأخرى بالصور ذاتها. إذن هناك بالفعل جوانبٌ عميقةٌ من التشابه تتصل

بالآلية التي ترى فيها هذه الشبكات الحاسوبية العالم.»
ويرى كلون أن ليس ثمة طريقة مُحكَّمة بعدُ من شأنها أن تحول دون انخداع الشبكات بأساليب خبيثة. ولكن ما إن تتحسن التقنية حتى «تُسدَّ الثغرات الأمنية، وتمسي الشبكات أكثر ذكاءً وفاعليةً. ويُنتظر أيضًا أن يكون أداؤها أفضل عندما تكون في مواجهة حالاتٍ مختلفةٍ جوهريًا عما تمرَّست عليه.»
واليوم تُدرَّبُ الإنسالاتُ وتُوطَّنُ -كما يقول كلون- على نوعٍ واحدٍ من الصُّور مستمدًّا من العالم الطبيعي، فإذا واجهتُ صُورًا متباينةً جدًّا تعطلَّتْ وبدت منها مظاهرُ سلوكٍ غريبةٍ وناشزة. «إنها بحاجةٌ فعلاً إلى أن تكون قادرةً على رؤية العالم وتمثُّل ما تراه في آنٍ معًا.

مراجع للاستزادة

- *Purves, D. and Lotto, R. B. Why We See What We Do: An Empirical Theory of Vision. 2011 ISBN-10: 0878935967*
- *Macknik, S.L. and Martinez-Conde, S. Sleights of Mind: What the Neuroscience of Magic Reveals About Our Everyday Deceptions. 2010 Henry Holt and Company, LLC. ISBN: 978-0-8050-9281-3*
- *Sugihara, K. (1986). Interpretation of Line Drawing. MIT Press, Cambridge. <http://www.evolvingai.org/publications>*
- *Nguyen, A., Yosinski, J., and Clune, J. (2015) Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images. In Computer Vision and Pattern Recognition (CVPR '15), IEEE, 2015.*

المؤلفة

إستر شاين كاتبة مستقلة في مجال التقنية والأعمال، تقيم في منطقة بوسطن.

مطبوعات الجمعية العلمية السورية للمعلوماتية

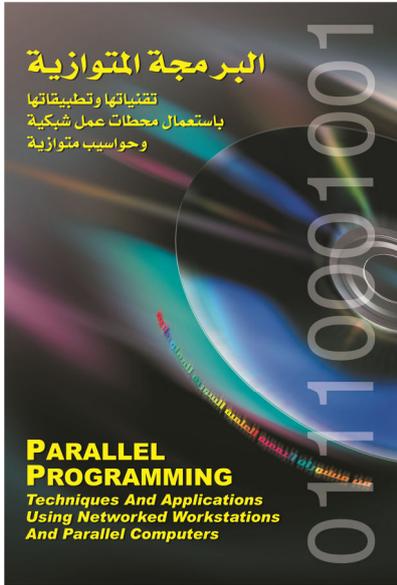
السعر (ل.س)			الكتب التخصّصية
مؤسسات	أعضاء جمعية وطلاب	أفراد	
4000	1600	2000	معجم مصطلحات المعلوماتية
2000	1200	1600	أسس لغات البرمجة
2400	1200	1800	هندسة البرمجيات - المجلد الأول
2000	800	1400	هندسة البرمجيات - المجلد الثاني
2000	1000	1500	الذكاء الصناعي
-	1000	1600	مفاهيم نظام التشغيل - الجزء الأول (تجليد عادي)
2200	1300	1900	مفاهيم نظام التشغيل - الجزء الأول (تجليد فني)
1900	1100	1600	مفاهيم نظام التشغيل - الجزء الثاني
3000	1600	2400	التعمية التطبيقية (Applied Cryptology)
-	400	600	المدخل إلى Mathematica 5.0 (تجليد عادي)
1200	600	800	المدخل إلى Mathematica 5.0 (تجليد فني)
1850	1100	1600	اتصالات المعطيات والحواسيب - الجزء الأول
1650	1000	1400	اتصالات المعطيات والحواسيب - الجزء الثاني
500	200	300	مسرد مصطلحات المعلوماتية إنكليزي - عربي
600	200	250	مجلة الثقافة المعلوماتية
2400	1200	1800	مدخل إلى الخوارزميات - الجزء الأول
2400	1200	1800	مدخل إلى الخوارزميات - الجزء الثاني
1800	1800	1800	أسس نظم قواعد المعطيات - الجزء الأول

صدر كتاب:

البرمجة المتوازية - تقنياتها وتطبيقاتها باستعمال محطات عمل شبكية وحواسيب متوازية

هو الكتاب العاشر ضمن سلسلة الكتب التي تصدرها الجمعية العلمية السورية للمعلوماتية بعد:

- "أسس لغات البرمجة" Essential of Programming Languages،
- "هندسة البرمجيات" Software Engineering - المجلد الأول والثاني،
- "الذكاء الصناعي" Artificial Intelligence،
- "مفاهيم نظام التشغيل" Operating System Concepts - الجزء الأول والثاني،
- "التعمية التطبيقية" Applied Cryptography،
- "المدخل إلى Mathematica 5.0"،
- "اتصالات المعطيات والحواسيب" Data and Computer Communications - الجزء الأول والثاني،
- "مدخل إلى الخوارزميات" Introduction to Algorithms - الجزء الأول والثاني،
- "أسس نظم قواعد المعطيات" Fundamentals of Database Systems - الجزء الأول.



وتُعد هذه الكتب كتبًا مرجعية لطلاب الهندسة المعلوماتية وهندسة الحواسيب وعلوم الحاسوب، وبذلك فإن الجمعية العلمية السورية للمعلوماتية تُسهم إسهامًا كبيرًا في سد النقص الواضح في المكتبة العربية التي تعاني فقرًا حقيقيًا في مثل هذه الكتب التخصصية.

الغرض من الكتاب تقديم تقنيات البرمجة المتوازية لحل مسألةٍ ما بسرعةٍ حسابية أعلى من سرعة حلّها باستخدام حاسوب منفرد.

قسّم الكتاب إلى بايين؛ يتناول الباب الأول التقنيات الأساسية للبرمجة المتوازية وتتضمن: مفاهيم الحواسيب المتوازية والحشود، ومساقات التراسل وبرمجيات MPI وتقويم أداء البرامج، والحسابات المتوازية طبيعيًا، والتجزئ وفرق تُسد، والتوارد، والحسابات المتزامنة واللامتزامنة، وموازنة الحمل، والبرمجة باستخدام الذاكرة التشاركية.

يناقش الباب الثاني خوارزميات ذات مسائل خصوصية في النطاقين العددي واللاعدي؛ كمسألة الفرز، والخوارزميات العددية، وضرب المصفوفات، والمعادلات الخطية، والمعادلات التفاضلية الجزئية، ومعالجة الصورة، والبحث والاستمثال.

ومن المزايا المهمة لهذا الكتاب أنه يعرض في نهاية كل فصل مجموعة كبيرة من المسائل المستنبطة من الحياة الواقعية، والمفيدة في تطوير المهارات في استخدام تقنيات البرمجة المتوازية.