"البحث العميق على الويب" ربما يفيد الباحثين.



توضح الصورة: ما تراه عندما تقوم ببحث على الويب هو رأس الجبل الثلجي فقط. أغلب المعلومات مدفونة في العمق. يقوم "مختبر الدفع النفاث" JPL و "وكالة مشاريع البحوث المتطورة الدفاعية" DARPA بالتعاون في مبادرة أطلق عليها اسم "ميميكس" Memex والتي تهدف إلى الكشف عن الروابط بين المعلومات الصغيرة المدفونة في محيط المعلومات الضخم.

عندما تقوم بعمل بحث بسيط عن موضوع على الويب فإن النتائج التي تظهر لا تعطي الصورة الكاملة. يحتوي الإنترنت على كم هائل من المعلومات، يطلق عليها أحيانًا "الويب العميق" **Deep web**، وهي البيانات التي لم يتم فهرستها من قبل محركات البحث، كالمعلومات التي من شأنها أن تكون مفيدة في تتبع المجرمين والأنشطة الإرهابية، أو الإتجار بالجنس وانتشار الأمراض. ويمكن للعلماء الإستفادة منها

في البحث عن صور وبيانات المركبات الفضائية.

تقوم وكالة مشاريع البحوث المتطورة الدفاعية **DARPA** بتطوير أدوات كجزء من مشروع ميميكس والتي تهدف إلى الوصول إلى هذا العالم الغامض على الإنترنت وفهرسته. وانضم باحثون من مختبر الدفع النفاث التابع لناسا في باسادينا ليساعدوا في البحث عن فوائد الويب العميق للعلم. على سبيل المثال يمكن لميمكس بأن تساعد على فهرسة الكم الهائل من المعلومات التي تنتجها المركبات الفضائية لناسا بشكل يومي.

ويذكر كريس ماتمان **Chris Mattmann** الباحث الأساسي من **JPL** في مشروع ميمكس: "نحن نقوم بتطوير الجيل التالي من تقنيات البحث والذي يفهم حاجات الأشخاص، الأماكن، الأشياء والروابط بينها."

لا يقتصر عمل ميميكس بالتحقق من المحتوى النصي فقط، بل يعمل أيضاً على التحقق من الصور ومقاطع الفيديو، والإعلانات، والبرامج النصية والعديد من الوسائل التي يتم فيها تخزين المعلومات؛ للنظر في كيفية ترابطها.

ويذكر كريس ماتمان: "نحاول أن نجعل من المفهرس التلقائي أو زاحف الشبكة (**Web crawlers**) أن يتصرف مثل المتصفحات في التعامل مع المعلومات. بعبارة أخرى: تنفيذ البرامج النصية وقراءة الإعلانات كما يفعل أي شخص على الإنترنت. وهذه هي المعلومات التي عادة لا يتم فهرستها في محركات البحث".

هذه الصورة ملتقطة من برنامجي ImageSpace و ImageCat في ميميكس والتي تظهر إمكانية أداة البحث من ترتيب كم هائل من المعلومات ومحاولة إيجاد روابط بين المعلومات المخبأه في الصور. حقوق الصورة: ناسا NASA/جي بي إل JPL–Caltech /كيت ويز Kitware/ كونتينيوم Continuum.

وبالإضافة إلى ذلك، البحث العادي على الويب لا يجلب معلومات دقيقة من الصور ومقاطع الفيديو، ولكن باستطاعة ميميكس أن يتعرف على محتوى تلك الصور ويربطها مع مواضيع البحث المناسبة. يمكن لأداة البحث أن تتعرف على نفس العنصر في إطارات مقطع فيديو أو حتى في مقاطع فيديو مختلفة.

إمكانية ميميكس في البحث في محتوى الصور ومقاطع الفيديو، يمكن أن تعود بفائدة للمهمات الفضائية التي تقوم بالتقاط صور ومقاطع فيديو وأنواع أخرى من تصوير المعلومات بأدوات مختلفة مثل جهاز المطياف المتري. كما أن البحث في المعلومات البصرية لجسم كوني معين سوف يسهل عمل العلماء كثيرًا في تحليل الخصائص الجيولوجية. ويمكن كذلك أن يستفيد العلماء من تحليل المعلومات البصرية التي تنتج من المهمات المتعلقة بالأرض والتي ترصد ظواهر مثل تساقط الثلوج ورطوبة التربة.

سيتمكن ميميكس من تعزيز البحث عن البيانات العلمية المنشورة، بحيث يمكن للعلماء أن يكونوا على دراية عن المعلومات المنشورة والتي تم تحليلها في نفس مجال بحثهم. ويمكن تطبيق هذه التقنية لمراكز وكالة ناسا الكبيرة مثل مركز الأرشيف النشط المنتشر لعلوم المحيطات الفيزيائية **Physical Oceanography Distributed Active Archive Center** والتي تساعد على جعل بيانات ناسا عن المحيطات والمناخ سهلة الوصول ومترابطة.

وسيجعل ميميكس ملفات البي دي إف **PDF** سهلة الوصول والبحث ومساعدة الأشخاص للوصول إلى المعلومة التي يريدونها بسهولة، كما أن الوعي بالمعلومات المنشورة والموجودة مسبقًا سوف يساعد مديري البرامج على تقييم أثر المركبات الفضائية.

كل النصوص البرمجية **codes** المكتوبة لمشروع ميمكس هي مفتوحة المصدر. وتمثل **JPL** أحد الفرق السبعة عشر التي تعمل في هذا المشروع مجزء من مبادرة داربا **DARPA**.

ميمكس هو جزء من مبادرة داربا للبيانات الضخمة **Big data** والتي أُطلق عليها **XDATA**، وقام بإدارته ويد شين **Wade Shen** مدير البرامج في داربا. ويهدف هذا المشروع البحثي لمعالجة وتحليل الكم الهائل من البيانات، المرتبطة بالتطبيقات الدفاعية والحكومية والمدنية. و تكون المشروع من 24 فريق مشارك أحدها هو **JPL**.

ويذكر ماتمان: "نقوم بتطوير منتجات مجانية ذات مصدر مفتوح ومروسة بعناية، ومن ثم تعزيزهم بواسطة استثمارات داربا ونقلها للمجتمع العلمي". وتعاونت شركة **Continuum Analytics** من أوستن في تكساس، وشركة كيت وير **Kitware Inc**. من كليفتون بارك في نيويورك، مع **JPL** في مشروع ميمكس. وتعتبر **JPL** جزء من معهد كاليفورنيا للتقنية.

- التاريخ: 2015-06-21
- التصنيف: الأرض

#الانترنت #المطياف المتري #ميميكس #Memex

المصطلحات

- **البيانات الضخمة (Big data)**: مصطلح متطور يصف كمية ضخمة جداً من البيانات المُهيكلة وغير المُهيكلة، يمكن تحليلها حسابياً للحصول على الأساليب والنزعات والعلاقات والروابط، خصوصاً تلك التي تتعلق بالسلوك والتفاعل البشري.

المصادر

- ناسا

المساهمون

- ترجمة
  - نداء البابطين
- مُراجعة
  - همام بيطار
- تحرير
  - هبة الأمين
- تصميم
  - نادر النوري
- نشر
  - مي الشاهد