

réponses sont possibles. Chacune a ses avantages et ses inconvénients.

Une première possibilité serait de sous-traiter la préservation des données à quelques grands opérateurs chargés de les recopier régulièrement sur les nouveaux supports avant que les anciens ne se dégradent et, éventuellement, de les convertir dans les nouveaux formats. À cet égard, les offres d'hébergement gratuit ne manquent pas, qu'il s'agisse d'abriter nos photographies, vidéos, documents bureautiques ou encore notre courrier. Le plus souvent, elles sont le fait de grands fournisseurs de services sur Internet, tels *Google, Amazon, Facebook, Apple, Microsoft*, etc. Toutefois, les récentes révélations sur le programme de surveillance PRISM de l'Agence américaine de sécurité soulignent un phénomène général : nous payons ces services en fournissant involontairement des informations très détaillées sur nos activités, préférences, relations personnelles, achats, voyages, etc.

Ces données sont employées à notre insu, ce qui se traduit non seulement par une surveillance occulte, mais aussi de plus en plus souvent par des violations à grande échelle de la vie privée, par exemple par des vendeurs de toutes sortes de produits.

D'autres grands acteurs ont lancé un gigantesque effort de préservation des données disponibles sur Internet ou sur certains réseaux sociaux : fondée en 1996, l'*Internet Archive* fait régulièrement des copies des sites disponibles sur la Toile et conserve une vaste collection d'images, de vidéos, de livres et de logiciels. S'agissant des œuvres littéraires, musicales ou cinématographiques, de grandes sociétés se réservent le droit exclusif de les archiver, et interdisent au public de les copier ou de les échanger ; en outre, elles mettent parfois en place des mesures de protection contre la copie qui entravent la conservation de ces données à long terme.

Aussi, tous ces acteurs, aussi puissants soient-ils, finissent-ils par constituer des points de fragilité comparables à la grande bibliothèque d'Alexandrie des temps hellénistiques. Malgré les investissements faits pour créer des centrales de serveurs plus sûres et mieux organisées, la concentration de grandes masses de données sous le contrôle d'une seule organisation est toujours source de fragilité. Et au-delà du problème technique de la sauvegarde des données contre les pertes accidentelles, il



5. L'AGENCE AMÉRICAINE DE SÉCURITÉ (NSA ou *National Security Agency*) exploite la concentration des données personnelles chez les grands fournisseurs de services sur Internet pour surveiller la population. Dans le projet PRISM, révélé par le lanceur d'alerte Edward Snowden, elle se fonde sur la loi américaine régulant les activités d'espionnage pour se fournir en données à décrypter chez les gros opérateurs, tels *Google, Amazon, Facebook*, etc.

■ L'AUTEUR



Roberto DI COSMO, professeur d'informatique à l'Université Paris-Diderot, est membre du Laboratoire Preuves, programmes et systèmes du CNRS.

s'agit aussi de trier les informations qui méritent d'être préservées et celles qui peuvent être effacées. Or le fait qu'une seule organisation, ayant sa logique et ses intérêts particuliers, en décide, introduit manifestement un énorme risque de destruction. Aurions-nous aujourd'hui des éditions complètes initiales des *Fleurs du mal* de Charles Baudelaire (1821-1867) ou de *Madame Bovary* de Gustave Flaubert (1821-1880), si, il y a plus d'un siècle, alors que ces chefs-d'œuvre étaient censurés, leur conservation avait été confiée à une organisation étatique centralisée ?

L'échange entre pairs

Une solution alternative de conservation des données de l'humanité à long terme provient du très grand nombre des ordinateurs personnels et de la croissance exponentielle de la capacité de stockage des mémoires disponibles pour le grand public. Le déploiement d'un système d'échange pair à pair (de l'anglais *peer to peer*) rend envisageable la transformation du parc mondial d'ordinateurs personnels en une gigantesque base de données distribuées, puis de laisser chacun y répliquer et y partager ses données.

Sous réserve que les ordinateurs personnels continuent à exister (!), cette solution aurait l'avantage de faire l'impasse sur une autorité centrale fixant seule ce qui mérite d'être conservé. Cela permettrait à

En 1085, Guillaume le conquérant ordonna un recensement général en Angleterre, qui devint ensuite la base indiscutable pour établir les taxes royales. Les Anglais nommèrent *Domesday Book* (le livre du jugement dernier) le volume où les résultats étaient consignés (ci-dessous à gauche). En 1986, la BBC décida de stocker les données du nouveau *Domesday* – un nouveau recensement et une enquête générale qu'elle avait organisés – sur des vidéodisques de type *LaserDisc*, support que l'on pensait alors pérenne (au centre). Toutefois, il se périma si vite que, 25 ans plus tard, il était déjà devenu impossible d'en lire le contenu ! Un long et coûteux effort technique fut alors nécessaire pour reproduire sur un ordinateur moderne le fonctionnement d'un lecteur de *LaserDisc*, ce qui conduisit au *Domesday reloaded* (*Domesday* refait), une nouvelle version consultable (à droite) qui, par prudence, a été versée aux Archives nationales britanniques...



l'humanité de décider spontanément des informations qu'elle souhaite conserver... Les années fastes du téléchargement de musique sur les anciens systèmes pair à pair nous ont donné un aperçu de ce qu'un tel réseau peut accomplir : à côté de tubes et autres succès commerciaux plus ou moins faciles, on y retrouvait souvent des morceaux de musique rares et, sinon, indisponibles. Il suffit en effet qu'existe un petit groupe de passionnés par tel ou tel type de données pour que celles-ci soient préservées quelque part dans le réseau social. Cela constitue un atout en faveur de la conservation de données à long terme, étant donné qu'il est difficile d'identifier quelle information pertinente aujourd'hui le sera demain ou, inversement, quelles données apparemment sans intérêt aujourd'hui pourraient se révéler précieuses dans l'avenir.

Comment résoudre les conflits d'intérêt ?

Encore faut-il que l'accès aux données dans la durée soit garanti par l'usage de formats ouverts et de logiciels libres. Dès lors, il serait utile de mettre en place une infrastructure de recherche répartie, comme les données, partout dans le monde, pour que les informations y soient stockées et indexées.

BIBLIOGRAPHIE

J. Rothenberg, *Avoiding technological quicksand : Finding a viable technical foundation for digital preservation*, rapport au *Council on Library and Information Resources*, janvier 1999 (www.clir.org/pubs/reports/rothenberg).

P. et S. Aigrain, *Sharing : Culture and the Economy in the Internet Age*, Amsterdam University Press, 2012.

Th. Priol, *Vers le tout-en-réseau ?*, Dossier *Pour la Science*, n° 66, janvier-mars 2010.

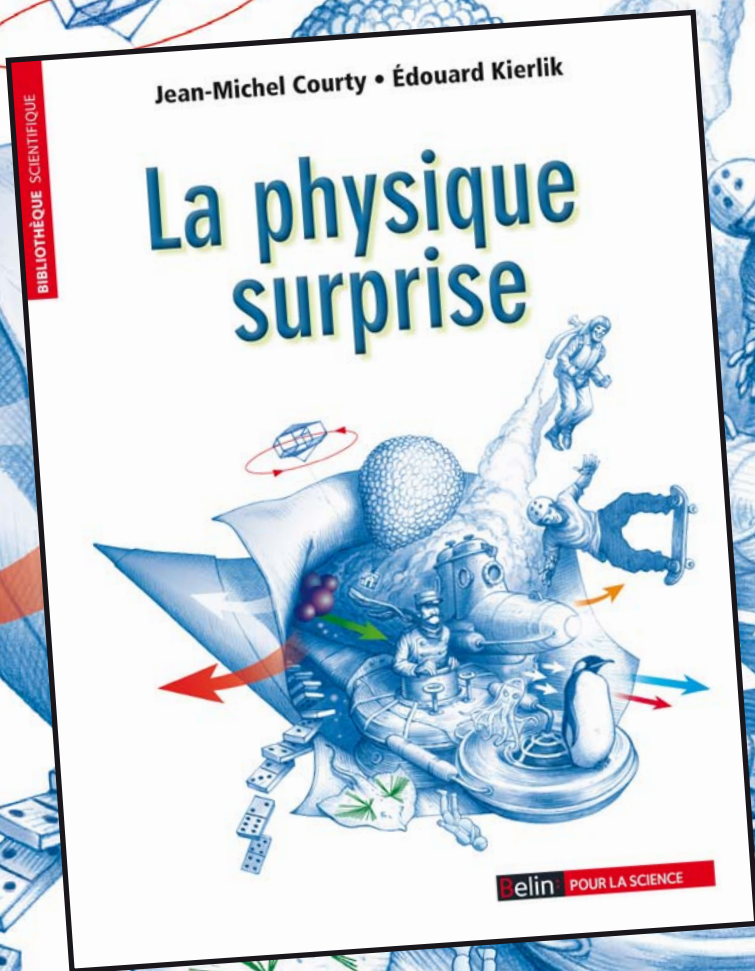
Sur le parc de PC prévisible à l'échelle mondiale : www.gartner.com/id=1602818

Le principal obstacle n'est pas technique, mais économique. Nombre d'acteurs des industries de la musique, du cinéma et du livre militent pour qu'on restreigne les échanges entre pairs et pénalise ceux qui y participent ; fondés sur la vente de copies individuelles, leurs modèles économiques s'écroulent si l'on donne la possibilité de réaliser à volonté un nombre illimité de copies numériques de toute œuvre. C'est pourquoi ils souhaitent l'application de lois contraignant les internautes à se détourner du modèle pair à pair pour passer à des échanges cryptés, lesquels ne peuvent que restreindre l'accès public aux données.

Nous ne pouvons détailler ici les propositions qui ont été faites pour concilier ces intérêts contradictoires, mais il nous paraît essentiel de souligner que le libre partage sans entrave est un instrument essentiel pour la préservation de nos données pour les futures générations. Il doit être encouragé et accompagné. De la même façon que les collections privées de livres ont préservé des titres dont les éditeurs ont disparu, les très nombreuses collections de copies numériques individuelles de données nouvelles ou anciennes seront pour nos descendants un moyen de conserver la culture actuelle, pourvu qu'elles soient enregistrées dans des formats ouverts, avec des logiciels libres.

La physique surprise

Jean-Michel Courty et Édouard Kierlik



Que nous réserve cette Physique surprise ? Trente-neuf pièces à déguster sans modération, mitonnées par les auteurs. Le style et les croquis, l'actualité et la variété des thèmes vous séduiront, vous amuseront, vous surprendront. Oui, on peut franchir le mur du son en chute libre. Non, les veines ne sont pas bleues. Oui, on peut séparer le chaud et le froid dans un courant d'air. Alors... Bon appétit !

Éditions Belin/Pour la Science 2013
184 pages – 24 euros
ISBN 978-2-8425-5121-9

BON DE COMMANDE

à découper ou à photocopier et à retourner accompagné de votre règlement à : Groupe Pour la Science • 628 avenue du Grain d'Or • 41350 Vineuil • Tél. : 0 805 655 255 • e-mail : pourlasciencevpc@daudin.fr

**POUR LA
SCIENCE**

☐ Oui, je commande l'ouvrage ***La physique surprise*** de J.-M. Courty et É. Kierlik.

..... exemplaires $\times$ 24,00 € (Réf. 075121) = €

Frais de port (4,90 € France – 12 € étranger) + €

TOTAL À REGLER = €

J'indique mes coordonnées :

Nom: _____

Prénom: _____

Adresse: _____

C.P.: [][][][][] Ville: _____

Pays: _____ Tél.: | | | | | | | | | |

E-mail pour recevoir la newsletter *Pour la Science* (à remplir en majuscules).

Je choisis mon mode de règlement :

☐ par chèque à l'ordre de *Pour la Science*

☐ par carte bancaire

Numéro

Date d'expiration

Cryptogramme

Signature obligatoire

En application de l'article 27 de la loi du 6 janvier 1978, les informations ci-dessus sont indispensables au traitement de votre commande. Elles peuvent donner lieu à l'exercice du droit d'accès et de rectification auprès du groupe Pour la Science. Par notre intermédiaire, vous pouvez être amené à recevoir des propositions d'organismes partenaires. En cas de refus de votre part, merci de cocher la case ci-contre ☐

PSPLS433

Offre valable jusqu'au 31.12.13, dans la limite des stocks disponibles.