

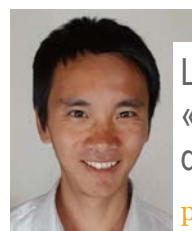
Marie Grégoire : « Une banque de données majeure pour le Québec »

p. 7



Guillaume Meyer, à l'écoute des chiffres

p. 43



Lê Nguyễn Hoàng : « Il faut restreindre les droits d'accès des agents IA »

p. 44

archimag.com

[STRATÉGIES & RESSOURCES DE LA MÉMOIRE & DU SAVOIR]

n° 397

21 euros - septembre 2026



[dossier]

documents, données, IA :
bâtir une chaîne
de preuve conforme



chaîne de confiance

La Cnil a reçu 20 150 plaintes en 2025 (soit 10 % de plus qu'en 2024), notamment pour non-respect de la protection des données et de la vie privée. 6 167 violations de données lui ont par ailleurs été notifiées. Parallèlement, les fraudes documentaires détectées dans le secteur de l'assurance ont bondi de 30 % pour atteindre un montant de plus de 900 millions d'euros. 13 % des Français reconnaissent avoir déjà falsifié un document, dont un tiers avec l'aide de l'IA.

Pendant longtemps, la gestion documentaire était considérée comme une affaire de classement, de stockage et d'archivage. Une facture arrivait, un contrat était signé, on numérisait un dossier RH, puis chacun espérait le retrouver le jour venu. Dans le même temps, la conformité était abordée comme un exercice passif, une case à cocher en fin de projet ou une contrainte administrative déléguée aux juristes. Ce temps est révolu. Un document numérisé, un RIB ou une facture PDF ne constituent plus, par leur seule présence, la garantie d'un fait réel. Le 1^{er} septembre 2026, toutes les entreprises devront recevoir des factures électroniques (les ETI et les grandes entreprises devront aussi les émettre)... L'AI Act déploie progressivement ses obligations, avec des exigences de supervision, de documentation,



Clémence Jost

DR

de transparence et de contrôle... Et le Cyber Resilience Act ajoute une pression nouvelle : celle, pour les fournisseurs de produits numériques, de documenter la sécurité et le suivi des vulnérabilités de leurs outils.

organiser une continuité

Chaque information traitée doit raconter son histoire : qui a reçu le document ? Dans quel format ? Quelles données ont été extraites ? Par quel outil ? Et avec quel degré de fiabilité ? Qui l'a modifié, validé ou archivé ? Des questions qui deviennent éminemment stratégiques dès qu'un litige, un audit, un contrôle ou une fuite

de données survient. Et la preuve réside moins dans le document final que dans les métadonnées, les journaux, les versions, les droits d'accès, les règles métier et les traces laissées par les systèmes.

Certes, les outils existent : capture intelligente, IDP, Ged, ECM, BPM, signature électronique, archivage probatoire... Mais leur empilement ne garantit rien. Mal paramétrés ou mal gouvernés, ils fabriquent de nouvelles zones grises : données incomplètes, durées de conservation incohérentes, traitements impossibles à expliquer. Notre dossier en témoigne : bâtir une chaîne de preuve conforme, ce n'est pas acheter une solution miracle. C'est organiser une continuité : de la capture à l'archive, de la donnée extraite à la décision tracée, de l'automatisation au contrôle. Si cette exigence peut paraître lourde, elle est surtout une chance. Celle de transformer la conformité subie en une méthode de fiabilisation documentaire, de réduction des risques et de confiance. ■

Clémence Jost

[Rédactrice en chef]

nous faisons Archimag

Serda édition-IDP
24, rue de Milan, F-75009 Paris
Tél. : +33 (0)1 44 53 45 00
Fax : +33 (0)1 44 53 45 01
infos@archimag.com
www.archimag.com
contacts e-mail
prenom.nom@archimag.com
rédaction
directeur de la publication
Pierre Fuzeau
directrice de la rédaction
Louise Guerre
rédactrice en chef
Clémence Jost
l'équipe de rédacteurs
Fabien Carré-Marillonnet,
Sivagami Casimir, Elisabeth
Hutin-Baillet, Éric Le Ven,
Bruno Texier

nos correspondants
Arbido (Genève), Ikram Bouz-
kar (Tanger), Marie-Laurence
Dubois (Belgique)
ont collaboré à ce numéro
Philippe Nieuwbourg,
Philippe Salaün, Cécile Touitou
site web, newsletter
Sivagami Casimir
conception graphique
Julio Arias-Aranz, Amcoot
maquette
Bruno Dalèle
brunodalele.com
dessinateur
Vince (vince-cartoon.be)

publicité
directrice de la publicité
Cathy Potel
01 44 53 45 00
chef de publicité
Imane Erraoui
01 44 53 45 00
responsable marketing
Albane Perrichon
vente au numéro
service abonnement
Zamila Nguyen
zamila.nguyen@archimag.com
réclamation et informations
commandes@archimag.com
24 rue de Milan
F-75009 Paris
tarifs et conditions
d'abonnement
valables jusqu'au 31/12/2026
France : 1 an, 189 euros
France : 2 ans, 338 euros

tarif étudiant : 1 an, 36 euros
tarif demandeur d'emploi : 1 an,
74 euros
vente au numéro : 21 euros
France : 1 an,
Pack abo : 402 euros
imprimeur
Inore Groupe Impression
4 rue Thomas Edison
58640 Varennes Vauzelles
éditeur
IDP Sarl, au capital
de 40 000 euros
Information,
documentation, presse
N° de commission paritaire :
0127 T 85484
ISSN : 2260-166X
Dépôt légal à parution
du numéro

crédits photos
Couverture : jubayerahmedr11/
Magnific
Intérieures : droits réservés,
sauf mentions différentes.
annonceurs
Calico France : 8, 9
IDP : encart, 2^e de couverture, 6,
4^e de couverture
Kodak : 19
Numen : 21
Open Source Experience : 2
Sage : 17
Serda : encart
Smart Tech Forum :
3^e de couverture
Tech Show Paris : 13
Les marques citées dans
le présent numéro sont
des marques déposées.

Accédez à nos réseaux
sociaux via ce QR code



Archimag est une publication
du groupe Serda. Toute
adaptation ou reproduction
même partielle des informations
parues dans Archimag est
formellement interdite sauf
accord écrit d'IDP SARL.
© IDP septembre 2026



Ce document est imprimé
sur papier certifié PEFC (respect
des fonctions environnementales,
économiques et sociales des forêts)

Pour vous abonner, voir page 47 ou sur www.archimag.com/boutique

Annoncez-vous sur Archimag et Archimag.com

Contactez Cathy Potel : cathy.potel@archimag.com,
Imane Erraoui : imane.erraoui@archimag.com et Angeline Piquery : angeline.piquery@archimag.com



sommaire

[actualités]

- 04 en forte hausse, la fraude documentaire se banalise
- 05 DocLang, le document pensé pour l'IA
- 06 une calculette carbone pour les bibliothèques
- 07 3 questions à Marie Grégoire (BAnQ)
- 10 DGFIP piratée : l'État sous pression
- 11 suite sécurisée pour la défense
- 11 maturité numérique : la France à la traîne
- 12 la start-up du mois : CarbonelO



04

[dossier]

- 14 **documents, données, IA :
bâtir une chaîne de preuve conforme**

La conformité documentaire change d'échelle. Entre la facturation électronique, le RGPD, Nis 2, l'AI Act et les exigences accrues de traçabilité, les organisations ne doivent plus seulement conserver des documents : elles doivent prouver l'origine des données, la fiabilité des traitements, la validité des décisions et la maîtrise des risques liés aux workflows, aux droits d'accès ou à l'IA.

Sommaire p. 15

[outils]

- 26 les Archives nationales de Monaco célèbrent leur premier anniversaire
- 28 fiche métier : être archiviste en 2026
- 30 les (nouvelles) pratiques des étudiants et l'évolution de leur rapport aux BU
- 32 Ori, le compagnon IA expert en orientation de l'Étudiant
- 34 l'IA s'impose dans les outils de datavisualisation
- 38 choisir un fournisseur d'IA : attention aux clauses contractuelles !
- 40 8 ans de protection des données : quel bilan pour les citoyens et les entreprises ?
- 42 favoris : trois outils pour créer des jeux de piste interactifs
- 43 Guillaume Meyer, à l'écoute des chiffres



38

[perspectives]

- 44 Lê Nguyễn Hoang : « Il faut restreindre les droits d'accès des agents IA »
- 46 Archimag store : applis, cadeaux, beaux livres, expos...
- 48 dans les archives d'Archimag, septembre 2006



46

DR

DocLang, le document pensé pour l'IA



« Le PDF a été conçu pour l'impression. DOCX a été conçu pour les éditeurs. DocLang est conçu pour ce qui vient ensuite... » : l'IA. IBM, Abbyy, Red Hat, HumanSignal, Nvidia et la Fondation Linux se sont regroupés pour développer DocLang, une norme documentaire lisible par l'IA. Ni convertisseur ni API, l'objectif de DocLang est de créer une norme « comme JSON pour les données, comme HTML pour le web ». ■

analyser les documents sans erreur

Selon ses promoteurs, DocLang est né d'un constat : les pipelines d'IA supposent des données propres et structurées. Pourtant, les documents réels (contrats, factures, documents réglementaires, etc.) ne répondent pas à ces critères, générant des erreurs et des dysfonctionnements dans leur analyse : tableaux transformés en texte brut, figures disparues, métadonnées supprimées... À ce jour, les solutions Docling et Abbyy FineReader Engine prennent déjà en charge la norme DocLang. ■

AI Overviews : Google défie les éditeurs

Après un lancement mondial commencé en mai 2024, Google a déployé AI Overviews (Aperçus IA) en France le 22 juillet dernier. Fondée sur Gemini, la fonction génère un résumé de réponses, dispensant les internautes de visiter les sites web. Une arrivée qui inquiète la presse française : si les éditeurs négocient avec Google leurs contrats au titre des droits voisins sur la réutilisation de leurs contenus depuis 2024, le géant du web a lancé son service sans attendre l'issue des discussions.

43 % de trafic en moins

Google assure laisser aux éditeurs le choix d'apparaître ou non dans ces résumés (opt-out). Pourtant, selon un rapport du Reuters Institute for the Study of Journalism de juin dernier, le trafic provenant des moteurs de recherche pourrait diminuer de 43 % au cours des trois prochaines années. L'Alliance de la presse d'information générale pose désormais la question du partage de la valeur créée à partir de ses contenus. ■

KB Crawl mise sur Mistral

KB Suite s'appuie désormais sur Mistral Small 3.1 pour assister les veilleurs dans leur phase d'analyse de l'information. Le modèle d'IA français interroge uniquement les bases documentaires des clients (RAG), et ceci sans que les données quittent l'environnement de KB Crawl (sur des serveurs dédiés). L'éditeur travaille par ailleurs à l'automatisation de l'analyse, de la classification, de l'enrichissement et de la diffusion. ■

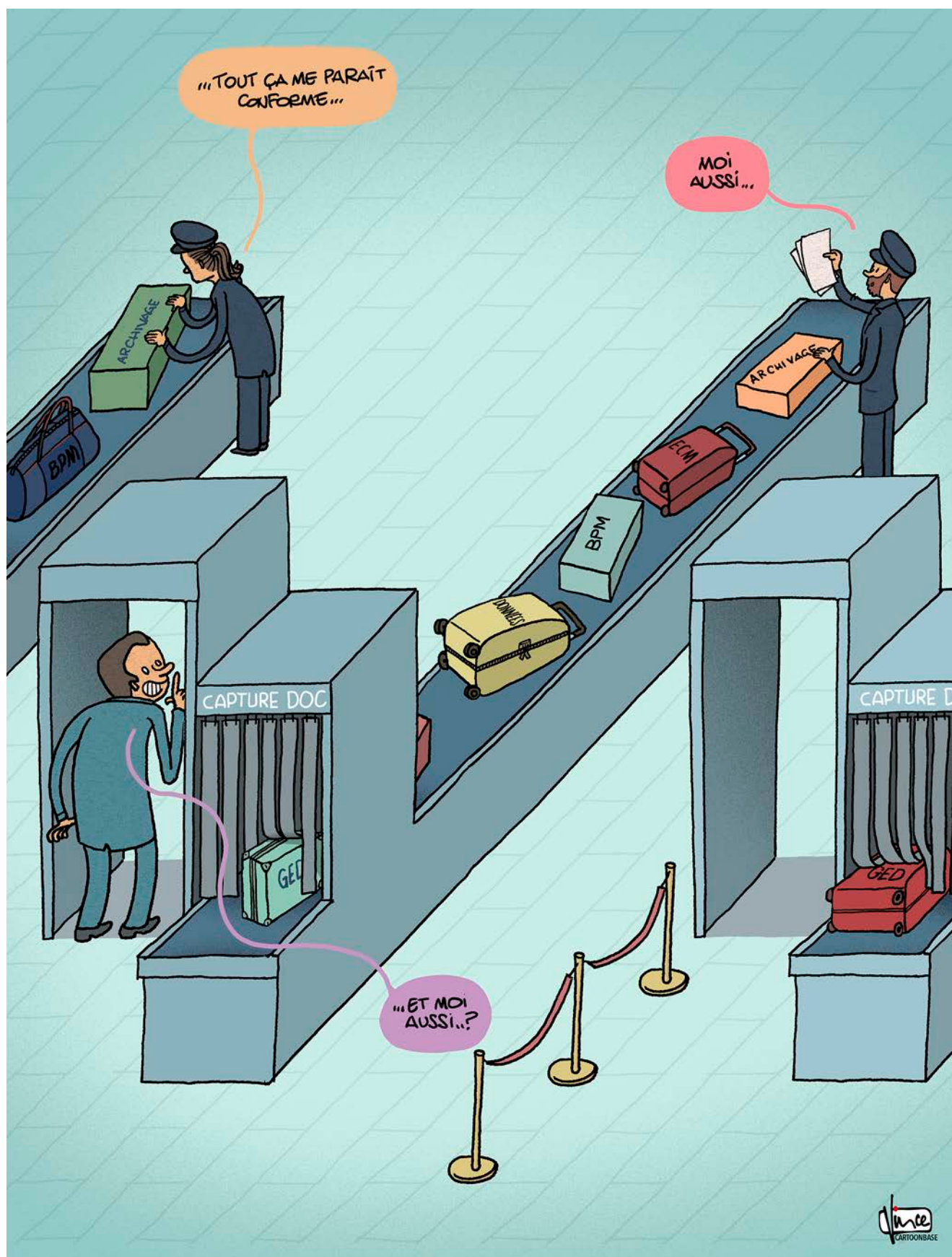
Mistral intègre Lexis+

LexisNexis ajoute les modèles Mistral aux autres IA déjà intégrées à sa plateforme juridique Lexis+ avec Protégé. L'éditeur entend ainsi garantir un cadre confidentiel, chiffré et hébergé en Europe pour des usages de rédaction, de synthèse ou de reformulation. LexisNexis précise que les prompts ne servent pas à entraîner les modèles d'IA et que les conversations sont supprimées après 90 jours. ■

données de la recherche

Le Catalogue national des données de Recherche Data Gouv a désormais son portail. Celui-ci centralise les jeux de données de la recherche scientifique française et en facilite l'accès grâce à un moteur de recherche renforcé par l'IA. Les utilisateurs peuvent effectuer leurs recherches en langage naturel et en plusieurs langues ou explorer les données par discipline. ■

→ catalogue.recherche.data.gouv.fr





15	documents, données, IA : bâtir une chaîne de preuve conforme
18	la capture, premier maillon de la preuve
20	Ged, ECM, BPM : passer du stockage au pilotage du risque
22	la conformité à l'épreuve de l'AI Act
23	de la donnée capturée à la preuve opposable
24	la fonction archive sous l'angle de la conformité
25	la checklist de la conformité

documents, données, IA : bâtir une chaîne de preuve conforme

La conformité documentaire change d'échelle. Entre la facturation électronique, le RGPD, Nis 2, l'AI Act et les exigences accrues de traçabilité, les organisations ne doivent plus seulement conserver des documents : elles doivent prouver l'origine des données, la fiabilité des traitements, la validité des décisions et la maîtrise des risques liés aux workflows, aux droits d'accès ou à l'IA. L'enjeu se déplace de la simple capture vers la gouvernance du cycle de vie de l'information. Une Ged, un ECM ou un BPM peuvent-ils vraiment piloter le risque ? Jusqu'où automatiser sans fragiliser la preuve ? Et sur quels critères choisir ses outils ? De la donnée extraite à l'archive opposable, chaque maillon engage aujourd'hui la responsabilité juridique, financière et réputationnelle de l'entreprise.

Le 23 septembre 1999, la sonde Mars Climate Orbiter de la Nasa se désintégrait dans l'atmosphère de Mars. En cause, un dysfonctionnement documentaire entre deux

logiciels : le premier communiquait la poussée des micropropulseurs en unités de mesure anglo-saxonnes, tandis que le second attendait des données en unités du système métrique. Cet accident

industriel s'est soldé par un double coût : une perte estimée à plus de 300 millions de dollars et des années de recherche réduites à néant.



agenda

Big Data & AI Paris**15-16 septembre 2026, Paris Expo, Porte de Versailles**

L'événement de référence du big data et de l'intelligence artificielle en France. Avec plus de 200 conférences, 220 entreprises exposantes et 17 000 participants.

→ www.bigdataparis.com**Salon SIDO 2026****16-17 septembre 2026, Lyon, Cité internationale**

Dédié à la transformation digitale des entreprises, le salon SIDO compte 70 conférences, 380 exposants et 210 speakers sur les innovations en matière d'IoT, d'IA, de XR et de robotique.

→ www.sido-lyon.com**Salon de la Data et de l'IA****22 septembre 2026, Nantes, Cité des Congrès**11^e édition de cet événement incontournable pour les professionnels de la data et de l'IA, organisé par l'IGN, Talan, KPC et Wavestone.→ salondata.fr**Congrès de l'ADBU****30 sept.-2 oct. 2026, Montpellier, Corum**Ce 55^e congrès aura pour thème « *Intégrités - (É)preuves et enjeux contemporains pour l'ESR* » et interrogera le positionnement et les missions des BU en faveur de l'intégrité sous toutes ses formes.→ adbu.fr/congres-2026**Smart Tech Forum****13 octobre 2026, Lille, EuraTechnologies**

Initié par l'Union des groupements d'achats publics (Ugap), cet événement rassemble les acteurs de la sphère publique autour des tendances et des dernières innovations technologiques.

→ www.ugap.fr/smart-tech-forum-c4608731au prochain
numéro

[dossier]

■ Patrimoine et musées : quelles innovations pour aujourd'hui et demain ?

[outils]

■ Fiche métier : qu'est-ce qu'un bibliothécaire en 2026 ?

■ Le labo, condition du métier de professionnel de l'info

■ IA dans les TPE/PME : la fuite invisible de la propriété intellectuelle

■ Comparatif des hébergeurs de données de santé (HDS)

■ Comment construire et faire vivre son référentiel d'archivage ?

dans les archives d'Archimag,
septembre 2006

au mois de septembre 2006, le Bureau central d'archives administratives militaires (BCAAM) de Pau se lançait dans un chantier titanesque : la numérisation de vingt millions de dossiers, représentant un total de 68 kilomètres linéaires. L'État y a mis les moyens : un parc informatique de 364 postes connectés sur un réseau par fibre optique, trois scanners tournant six heures par jour et deux baies de stockage en miroir sur deux sites distants. Il fallait bien ce déploiement de moyens pour traiter les millions de feuillets nominatifs de contrôle (FNC), qui relatent la carrière des administrés et qui étaient estimés à eux seuls à quinze téraoctets de données. D'autant que le BCAAM traitait alors par ailleurs, chaque année, 240 000 courriers entrants, principalement des demandes d'état signalétique et de services.

Qui dit chantier de numérisation dit préparation des documents. Il a en effet fallu nettoyer les documents, les dégrader et saisir une série de codes, avant de procéder à la numérisation elle-même. Heureusement, la technologie OCR (reconnaissance optique de caractères) prenait le relais pour analyser le document recto verso. « *Tout projet de ce genre est un projet de conduite au changement* », estimait le colonel Alain David, alors directeur du BCAAM. « *Il est passionnant dans une carrière de pouvoir participer à la mise en œuvre d'une telle ambition. Je savais commander des hommes. Il m'a fallu apprendre à dompter des machines* ».

imaginer les services
de demain

Notre édition du mois de septembre 2006 consacrait un autre reportage aux Jardins de l'innovation de France Télécom, ouverts depuis 2003 à



Couverture d'Archimag, septembre 2006. DR

Issy-les-Moulineaux (Hauts-de-Seine). En réalité, ces laboratoires de recherche et développement ont été répliqués ailleurs en France et regroupaient alors plus de 4 000 chercheurs issus de corps de métiers différents : mathématiciens, ingénieurs, sociologues, ergonomes... La mission qui leur était confiée consistait à imaginer les services de demain, comme ce « *mur de téléprésence* », qui donnait l'impression d'être en face de son interlocuteur, alors que celui-ci se trouvait à l'autre bout du monde. Ou bien ces prémices d'outils de visioconférences qui, vingt ans plus tard, sont devenus le quotidien de nombreux professionnels. Comme le confiait un ingénieur à propos des innovations qui étaient conçues dans ce laboratoire, « *l'usage que feront les utilisateurs nous échappe totalement. On peut s'attendre à des surprises !* » ■

Bruno Texier

Retrouvez la collection Archimag 1985-2015, 290 numéros en texte intégral sur :

→ collection.archimag.com