Proposition de messages à intégrer aux formations des usagers

= les données à caractère personnel (DCP) ont de la valeur et doivent être protégées
= **votre établissement n'est pas un sanctuaire**
= tous les éditeurs ne sont pas des courtiers de données (databrokers)
= enjeux à plusieurs échelles : individuelle, institutionnelle, nationale, européenne, internationale
= quelques **bonnes pratiques :**
- **limiter la quantité d'informations exposées** : se créer un compte individuel sur les plateformes docelec uniquement si vrai besoin, ne pas partager de donnée sensibles sur les forums et commentaires publics, utiliser des pseudo quand c'est possible.
- **utiliser les bons outils** : navigateurs axés sur la vie privée, extensions anti-traqueur, adresses e-mail secondaires temporaires, VPN. Refuser les cookies, limiter les autorisations des appli
- **nettoyer régulièrement ses traces** : vider régulièrement le cache de son navigateur, ne pas hésiter à faire valoir ses droits d'accès et d'effacement, certains sites spécialisés identifient les courtiers en données et expliquent comment demander le retrait de ses informations : JustDeleteMe (manuellement) ou Incogni (service payant)

Commission Pédagogie et Documentation

couperin.org

Groupe de Travail GPS / **Groupe Pistage et Surveillance** (créé en mars 2025 ; 5 personnes).
Objectifs généraux :
- faire l'audit de plateformes et de contrats (Dalloz, Cairn, Europresse), en coordination avec d'autres consortia étrangers (à l'image de ce que SPARC US a fait pour Elsevier et Springer-Nature)
- former et outiller les négociateurs Couperin pour mieux négocier les contrats et obtenir des conditions plus favorables que celles imposées par la loi
- promouvoir de bonnes pratiques/bons outils, à destination des bibliothèques et des utilisateurs de docelec
- alerter hors des établissements (CNIL, SupDPO, ministère, députés)

En Californie :
à partir du 1er août 2026, les Californiens pourront demander en une seule fois la suppression de toutes leurs données personnelles détenues par les courtiers en données,
en vertu du California Consumer Privacy Act (CPPA), sans avoir besoin de contacter les entreprises une à une comme en Europe avec le RGPD.
Lorsqu'un data broker recevra une demande d'effacement, il devra le faire "au moins une fois tous les 45 jours"
et ne pourra plus "vendre ou partager" de nouvelles données personnelles du requérant.
Tous les trois ans, les courtiers en données devront se soumettre à un audit mené par un tiers indépendant pour déterminer la conformité aux dispositions légales, avec astreinte financière de 200$ par jour.

CNIL. COMMISSION NATIONALE INFORMATIQUE & LIBERTÉS

Une réglementation protectrice en France et en Europe

Le RGPD limite le pistage et protège l'usager
- impose : licéité, transparence sur les moyens et les finalités, minimalisation des volumes et des durées, consentement et droit de contrôle…
- responsabilité des fournisseurs, des établissements et des utilisateurs finaux

Les délégués à la protection des données (DPO) en établissement et chez les éditeurs
accompagnent les services d'une organisation pour toutes les problématiques liées à la collecte, à l'utilisation et à la conservation de données personnelles.

Mais une réalité plus complexe :
- consentement réel ou implicite. Imposé ?
- information disponible, lisible, lue, comprise ?
- sous-traitance et transferts internationaux ?
+ *petits* éditeurs :
=> améliorations possibles quand prise de conscience
+ *gros* éditeurs :
=> négociations possibles selon le rapport de forces

Le pistage des usagers de la docelec

Les pratiques de pistage des données personnelles par certains éditeurs réinterrogent les politiques de formation des usagers.

L'usage de la documentation électronique est à la fois très onéreux et indispensable pour l'enseignement et la recherche.

Si c'est gratuit, c'est vous le produit,
Si c'est payant… aussi !

La docelec : un vrai besoin et des usages massifs
- 166 millions de téléchargements docelec en France selon l'ESGBU 2024,
- à comparer aux 6 millions prêts physiques

+ Un marché spécifique
Un article scientifique est singulier et non-substituable, chaque éditeur détient le monopole sur ses articles

Des milliers de petits éditeurs pas ou peu rentables + **Un oligopole de géants**
- cotés en bourse
- des fonds d'investissements parmi leurs actionnaires

position ultra-dominante
Elsevier publie 1 article scientifique / 5 dans le monde, toute discipline confondue

75% des budgets documentation dans l'ESR

nécessité de rentabiliser l'investissement par la formation des usagers

insoutenabilité financière

Impose des tarifs en augmentation constante, avec un modèle économique qui nous fait payer 3 fois :
- le salaire des chercheurs (écriture/relecture),
- les frais de publication (en OA ou non),
- les abonnements des bibliothèques

La rentabilité du secteur est supérieure à celle de l'armement, du pétrole, du luxe ou des GAFAM :
35 à 42% de bénéfices nets après impôts

Réaction de la communauté scientifique

développement de l'Open Science + renforcement du rôle des consortia : pour négocier les contrats, jusqu'à organiser des boycott internationaux
- Couperin et l'Aciege en France
- DEAL en Allemagne, SURF aux Pays-Bas…

le nouveau relai de croissance : les données personnelles

depuis 2024, Elsevier/RLEX se présente désormais à ses actionnaires comme *global provider of information-based analytics and decision tools* .
Le Chiffre d'Affaires de la partie courtage en données a dépassé celle de l'édition scientifique

Argumentaire des éditeurs :
pistage pour faciliter les usages, améliorer le service et en créer de nouveaux (authentification automatique, personnalisation, recommandations, etc.)

D'autres solutions garantissant l'anonymat existent

Des risques réels

1) **ciblage marketing** (intérêt, profilage du consentement à payer maximal) ex : majoration du tarif des billets d'avion
2) **discrimination** commerciale ou à l'embauche, sur critères sociaux, politiques ou médicaux ex : exclusion d'assurance/banque
3) **manipulation** électorale ou politique ex : scandale Cambridge-Analytica
4) risques liés à la **sécurité des données** (escroqueries et usurpation d'identité par des hackers…) https://bonjourlafuite.eu.org/
5) les **dérives autoritaires** (surveillance, atteinte à la liberté académique, liberté d'expression) et ses conséquences

L'ICE, le service de l'immigration des Etats-Unis, a signé en 2021 un contrat de 22,1M $ avec l'éditeur Lexis (base de presse Lexis Uni et base juridique Lexis360) pour l'accès à *Accurint*, un vaste référentiel d'informations publiques et privées, créé pour identifier rapidement des individus, les relier à leurs proches, à des lieux spécifiques, à leurs employeurs et à tout bien ou actif qu'ils pourraient posséder (entreprises, véhicules, etc).

La collecte des données personnelles

Exploitations légitimes + exploitations discutables | **Données académiques + données publiques + données privées** | **… par des moyens techniques variées**



le marché mondial du courtage de données représente environ :
- 270 Mds $ en 2024
- 473,35 Mds $ d'ici 2032
source : *Fortune Business Inside*

Finalités du pistage
- collecte de **données techniques** (fonctionnement, détection des usages illicites) ou de statistiques d'usages (notamment pour les bibliothèques)
- **amélioration de l'expérience utilisateur**, personnalisation des contenus
- **objectifs commerciaux** : segmentation des utilisateurs, renforcement du pouvoir de négociation commerciale, développement de services propriétaires (profils chercheurs, recommandations)
- **revente à des tiers**

= vraie valeur stratégique et économique

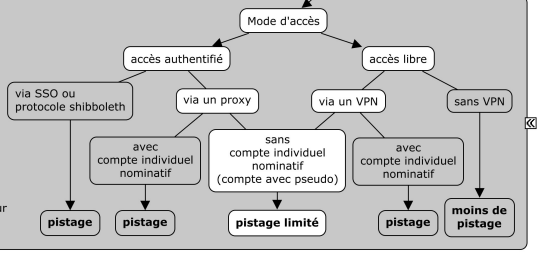
Détail des données potentiellement pistées
1) usages académiques :
- recherches effectuées et les contenus affichés ou téléchargés (quoi, quand, où)
- navigation (pages consultées, durée, vitesse de frappe, parcours de la souris…)
- le matériel utilisé (marque modèle, résolution de l'écran, polices installées, langue, fuseau horaire…)
- adresse IP, identifiants, catégories démographiques (nom complet date de naissance, genre…)

2) croisement possible avec les données publiques :
- actes de naissance/mariage, inscription listes électorales, titres de propriété…

3) croisement possible avec les données de vos usages perso :
- tél, adresse physique et mail, occupation et niveau de revenus, niveau d'éducation
- les interactions numériques : réseaux sociaux, prompts IA, historique et habitudes d'achat, services d'abonnements, programmes fidélité, enregistrement de garantie, parfois les journaux d'appels sur tél selon les permissions paramétrées
- appli installées, cookies déjà présents sur les navigateurs,
- localisation GPS et historique des déplacements
- données médicales (montre connectée…),
- mode de vie et intérêts : préférences et aversions, affiliations politiques, croyances religieuses, adhésions et affiliations à diverses organisations, loisirs et intérêts

Outils de pistage
- Cookies informatiques : fonctionnels, analytiques, ou tiers
- Identifiants persistants : compte utilisateur, SSO, adresse IP
- Pixels de suivi ("pixels espions")
- Empreintes numériques du matériel :
- Scripts externes (ex : Google Analytics)

Destination des données pistées
- transfert à l'éditeur
- ou stockage dans un 'espace local' de votre ordinateur, privatisé par l'éditeur



Couperin : Groupe Pistage et Surveillance