Dans le courant de l'année 2007, un agent public a envoyé un mail à un de ses homologues d'un pays sensible, sans se rendre compte qu'il avait associé au message quelques documents ultra confidentiels en pièce-jointe. Cet exemple confirme deux tendances que Benoit Dupont dessine dans son étude : l'administration publique est moins bien protégée que les entreprises tandis que la disparition d'informations dans les services gouvernementaux est résulte davantage de l'erreur humaine que du piratage. L'auteur, en procédant à une analyse sectorielle du phénomène, nous révèle que la disparition de données personnelles touche de manière très différente chaque secteur d'activité. Il en conclut que chaque organisation doit procéder à une analyse précise de sa situation et suggère un éventail de techniques de prévention devant permettre aux entreprises de mieux sécuriser leurs données personnelles.

Les organisations : sentinelles aveugles de la sécurité des données personnelles?

Benoît Dupont[1]
Chaire de recherche du Canada en sécurité, identité et technologie
Université de Montréal

Aux États-Unis, de 2006 à 2008, plusieurs grandes entreprises de distribution et des opérateurs de paiement se faisaient dérober plus de 200 millions de numéros de cartes de crédit par le pirate informatique Albert Gonzalez. En octobre 2007, le Ministère anglais du budget égarait les informations personnelles[2] et bancaires de 25 millions de bénéficiaires des allocations familiales. Ces deux incidents, abondamment médiatisés, ont permis aux opinions publiques occidentales de prendre conscience, d'une part de l'ampleur des quantités de données accumulées par les organisations publiques et privées sur leurs usagers ou leurs clients, et d'autre part des pratiques frisant l'insouciance qui régissent le traitement de celles-ci.

En effet, dans le premier cas, les mesures de sécurité déployées par certaines entreprises ciblées violaient impunément les normes de protection des transactions financières : certaines communications se faisant sur des réseaux sans fil n'étaient, par exemple, pas chiffrées ou utilisaient des solutions de chiffrement anciennes faciles à contourner en quelques minutes à l'aide d'outils disponibles sur Internet. Dans le second cas, les coûts induits par l'aseptisation des données (environ 15 000 livres ou 16 000 euros) furent jugés excessifs par l'employé responsable de l'opération.

Si ces deux affaires illustrent de manière particulièrement édifiante les conséquences catastrophiques de dysfonctionnements organisationnels en matière de protection des

[1] benoit.dupont@umontreal.ca

[2] On entend ici par 'informations personnelles' ou 'données personnelles' des informations ou des données qui permettent d'identifier un individu (son nom, son âge, son adresse, son numéro de sécurité sociale,…), de connaître ses capacités et habitudes de consommation (états des actifs financiers, biens et services achetés, modes de paiement,…), son statut à l'égard de certains services publics (dossier fiscal, dossier médical, éligibilité à des programmes sociaux, casier judiciaire, …), ou encore ses préférences personnelles (composition du réseau relationnel d'amis, goûts musicaux, …). Toutes ces données sont aujourd'hui recueillies et analysées de manière routinière et automatisée par la plupart des organisations publiques et privées.

données personnelles et de sécurité de l'information, elles ne constituent que des cas d'espèce d'une longue litanie d'incidents qui viennent faire, à intervalles réguliers, la une des médias. Il est néanmoins possible de dépasser cette dimension anecdotique pour analyser de manière plus systématique les facteurs de risques associés aux pertes et aux vols de données dans les organisations. Un tel exercice vise notamment à évaluer la performance des organisations publiques et privées en termes de sécurité des informations personnelles de leurs usagers, clients et employés, et à concevoir des stratégies de protection fondées sur des données probantes plutôt que sur de gros titres sensationnalistes ou des arguments commerciaux.
Dans une première partie, nous exposerons donc les résultats tirés de l'analyse des brèches informatiques survenues en Amérique du Nord entre 2005 et 2007, avant d'examiner, dans une seconde partie, comment de telles données pourraient être mises à profit afin d'élaborer des stratégies de prévention efficaces, ancrées dans l'approche criminologique de la prévention situationnelle[3].

1) Le profil des risques associés aux pertes et aux vols de données personnelles

Afin d'établir un profil des risques basé sur une quantité suffisamment représentative d'incidents, une base de données fut constituée. Elle comprend 976 dossiers de pertes ou de vols de données, survenus entre 2005 et 2007, impliquant 313 millions de dossiers personnels (Dupont et Gagnon 2008). Ces informations relèvent, aux Etats-Unis, d'obligations législatives de divulgation et de notification des victimes, auxquels sont soumises les entreprises et les services publics dans plus de 40 Etats, en cas de perte ou de vol de données.

Trois sources d'information existantes ont été consultées pour identifier les événements pertinents : la *Privacy Rights Clearinghouse*[4], le *Wayne Madsen Report*[5], et la liste élaborée par Rita Tehan (2007) pour le *Congressional Research Service*[6]. Pour chacun des incidents recensés, des recherches additionnelles furent conduites à l'aide des bases de données médiatiques afin de préciser les circonstances de l'incident et éviter le biais d'un nombre de dossiers personnels compromis surévalué[7]. Pour le Canada, 23 incidents associés à la perte ou au vol de 4,4 millions de dossiers ont été recensés, mais ces chiffres sous-estiment une réalité bien plus préoccupante, en l'absence d'un cadre réglementaire forçant les organisations à la transparence. Les informations recueillies sur chaque incident furent ensuite intégrées dans une base de données gérée à l'aide du

[3] Selon la définition classique de Maurice Cusson (2002, p. 39), « on entend par 'prévention situationnelle' les modifications des circonstances particulières dans lesquelles les délits pourraient être commis afin qu'ils paraissent difficiles, risqués ou inintéressants pour qui serait tenté de les commettre ».
[4] http://www.privacyrights.org/ar/ChronDataBreaches.htm#CP, consultée le 9 janvier 2010.
[5] http://www.waynemadsenreport.com/, consultée le 9 janvier 2010. Il est à noter que l'accès à cette page était gratuit jusqu'au début de l'année 2008.
[6] Depuis la fin de notre projet, une base de données interactive s'est ajoutée au trois sources principales que nous avons utilisées. Il s'agit de la Dataloss DB [database] de l'Open Security Foundation, accessible à http://datalossdb.org/, consultée le 9 janvier 2010.
[7] En cas d'estimations contradictoires, les chiffres les plus bas furent systématiquement retenus.

logiciel d'analyse statistique SPSS (*Statistical Package for the Social Sciences*) en fonction de variables telles que le nom de l'organisation concernée, la date de l'événement, le secteur d'activité de l'organisation, le type d'incident et ses causes, ou encore le nombre et le profil des victimes. Cinq grandes tendances se dégagent de ces analyses.

a. Un secteur public tout aussi exposé que le secteur privé

Tout d'abord, les services publics ne semblent pas en mesure de mieux protéger les informations personnelles de leurs usagers que les entreprises, puisque les secteurs de l'éducation, des services gouvernementaux et de la santé représentent conjointement 61,7% des incidents analysés[8], alors que le secteur financier n'arrive qu'en quatrième position avec 15,6%, suivi par le secteur manufacturier (7%) et le commerce de détail (6,9%). Ce constat est particulièrement préoccupant dans le contexte actuel de développement des services gouvernementaux en ligne, où l'efficacité recherchée ne semble pas accompagnée des capacités requises en termes de sécurité. Il semble en effet que les administrations n'aient pas encore pris conscience de la valeur, pour les fraudeurs, des gisements de données personnelles qu'elles exploitent.

b. La négligence, l'incompétence et l'indifférence bien plus préoccupantes que le piratage

L'analyse des causes à l'origine des incidents laisse entrevoir que le piratage est loin de représenter la principale menace, puisqu'avec 22,7% des incidents, il ne vient qu'en troisième position derrière la disparition d'équipement (40,1%) et la négligence ou l'erreur humaine (24,7%). La disparition d'équipement fait ici principalement référence au vol d'ordinateurs portables ou d'équipements de stockage des données (clés USB, disques durs amovibles, CD, etc.) par de petits délinquants qui sont seulement intéressés par la valeur de revente du matériel. De nombreuses pertes sont également inclues dans cette catégorie. En effet, d'après une étude du cabinet *Ponemon* (2008), on retrouvait dans les 36 plus gros aéroports américains en moyenne 10 000 ordinateurs portables « oubliés » chaque semaine par des voyageurs d'affaires distraits. Les négligences et les erreurs humaines concernent, quant à elles, de mauvais paramétrages techniques d'équipements ou d'applications, et reflètent la complexité croissante des systèmes de gestion des données personnelles, ainsi que le manque de formation de la plupart des opérateurs dans ce domaine. Malgré un discours médiatique mettant l'accent sur les exploits de pirates informatiques basés dans des pays émergents, force est de constater que la malveillance génère statistiquement moins de risques que l'indifférence des gestionnaires garants de la sécurité des informations personnelles.

L'origine des risques varie évidemment d'un secteur à un autre. Ainsi, pour le commerce de détail, c'est le piratage qui constitue la principale menace (37,3% des incidents), alors

[8] Les proportions restent comparables si l'on prend le nombre de dossiers compromis comme unité de référence, plutôt que le nombre d'incidents.

que le secteur financier est d'abord confronté au vol de ses équipements informatiques portables (46,1% des incidents dans ce secteur). Par ailleurs, une analyse transversale fait ressortir que la fréquence des incidents causés par la négligence est la plus basse dans le secteur manufacturier (15,9% des incidents) et la plus élevée au sein des services gouvernementaux (33,5% des incidents), ce qui pourrait s'expliquer par des différences marquées en matière de niveaux d'expertise.

c. Un volume moyen limité de dossiers compromis, mais de fortes variations sectorielles et techniques

Le nombre médian de dossiers compromis par incident est de 6 000, mais nous observons une forte amplitude dans l'échantillon où plusieurs cas de piratages impliquent des dizaines de millions de dossiers volés, souvent dans le secteur financier. En effet, des incidents d'origine criminelle vont générer en moyenne sept fois plus de dossiers compromis (847 000) que la négligence (129 000), et quatre fois plus que le vol d'équipement (211 000). Dans chaque secteur d'activité, le nombre moyen de dossiers compromis varie également considérablement, de 35 000 dans le secteur de l'éducation à un peu plus de deux millions dans celui du commerce de détail. Ainsi, la gravité des risques encourus par de tels incidents varie fortement en fonction de la taille des bases de données gérées, mais également de l'origine des incidents et certainement des compétences dont dispose chaque secteur d'activité pour en prévenir l'occurrence et en minimiser les conséquences.

d. Des victimes égales devant les risques

Une analyse de la répartition des victimes individuelles (c'est-à-dire les personnes physiques correspondant aux dossiers compromis) par catégorie démontre que les organisations éprouvent autant de difficultés à sécuriser les informations de leurs usagers[9] (44,3% des victimes) ou de leurs employés (22,8%) que celles de leurs clients (20,9%). Cette distribution équilibrée semblent indiquer que ce sont bien les modes actuels de gestion des informations personnelles qui posent problème aux organisations, quelles que soient les personnes concernées.

e. Une évolution des risques difficile à mesurer

Enfin, la nature récente des efforts de recueil des données sur ces incidents, ainsi que l'hétérogénéité de l'encadrement règlementaire entourant leur divulgation empêchent de mesurer l'évolution du phénomène avec un degré de confiance suffisant. En effet, le quadruplement des incidents enregistrés entre 2005 et 2006 doit-être davantage attribué à l'introduction de nouvelles obligations législatives qu'à une explosion du

[9] La distinction entre client et usager correspond à la nature payante ou non des services ou des biens offerts par l'organisation impliquée. Ainsi, les clients sont principalement concernés dans les incidents qui impliquent des institutions financières ou des entreprises du commerce de détail, alors que les usagers pourront aussi bien être des patients du système de santé que des étudiants ou des contribuables.

nombre des pertes ou des vols de données personnelles. De la même façon, la réduction significative observée en 2009 par l'*Open Security Foundation*[10] (-40% d'incidents par rapport à l'année précédente) reflète-t-elle une meilleure sécurité? Les usagers devraient s'en réjouir, mais elle traduit surtout un désintérêt des grands médias à rapporter ce genre d'événements dans un contexte de crise économique fournissant de nombreux autres sujets d'inquiétudes. Enfin, la diminution du nombre d'organisations remplissant leurs obligations de divulgation doit être également envisagée. Pourtant, des indicateurs fiables constituent des éléments indispensables pour l'évaluation de l'efficacité des stratégies de prévention et de contrôle déployées par les organisations.

2) Mesurer pour mieux prévenir : aligner les techniques de prévention situationnelle sur l'exposition aux risques

Plusieurs stratégies s'offrent aux organisations pour maîtriser ces risques omniprésents, de l'inaction à l'assurance ou au transfert à des tierces parties en passant par l'abandon pur et simple (Button 2009). C'est toutefois à la réduction des risques que nous consacrerons le reste de cet article, en préconisant une approche en trois étapes relevant :

1) de la connaissance des risques spécifiques auxquels sont confrontés une organisation et le secteur auquel elle appartient;

2) de la connaissance des techniques de prévention les plus appropriées pour y faire face;

3) de la connaissance des méta-risques, c'est-à-dire des risques découlant des activités de gestion des risques.

a. L'insécurité différentielle des données personnelles

Comme nous l'avons vu, chaque secteur d'activité se trouve confronté à des risques variant en fonction de leur probabilité, de leur gravité et de leur origine. Il devient alors possible de cartographier l'exposition au risque de chaque organisation par référence à son secteur d'appartenance. En ce qui concerne les données à notre disposition, la probabilité est déterminée par la proportion d'incidents observés dans chaque secteur, et la sévérité par le nombre moyen de dossiers compromis par chaque secteur. Nous obtenons alors un quadrant (Figure 1) dans lequel les organisations du secteur financier et les services gouvernementaux doivent faire face à des risques à la fois fréquents et sévères, alors que d'autres organisations doivent gérer des événements tout aussi graves mais plus rares (commerce de détail, industrie manufacturière) ou au contraire, potentiellement beaucoup plus fréquents mais aux conséquences plus limitées (santé, éducation)[11]. La détermination par une organisation de sa position dans l'un des secteurs du quadrant plutôt que dans un autre influencera notamment la nature et

[10] http://datalossdb.org/statistics.

[11] La tolérance au risque représente une troisième dimension qu'il ne nous a pas été possible de mesurer ici en raison de la nature des données utilisées.

l'ampleur des investissements de sécurité qui seront consentis. Par exemple, les organisations du secteur de l'éducation devraient mettre l'accent sur la formation de leurs employés, qui seront confrontés plus fréquemment aux incidents, alors que des secteurs où la sévérité prend le pas sur la probabilité auront intérêt à privilégier des investissements technologiques qui dépendent, dans une moindre mesure, de la vigilance des individus.

Bien que cette démarche d'évaluation semble élémentaire, il est surprenant de constater que la majorité des organisations connaissent mal les risques liés à la gestion de leurs données personnelles. Dans un sondage mondial, mené en 2008 auprès de 7 000 cadres, la firme de consultants PriceWaterhouseCoopers affirmait que 65% des répondants étaient incapables d'indiquer le nombre d'incidents de sécurité informatique subis par leur organisation au cours des 12 derniers mois, et 56% ne connaissaient pas la nature des incidents en question. Les connaissances relatives aux données personnelles gérées sont elles-mêmes très parcellaires. La même étude faisait ainsi ressortir que seulement 29% des organisations disposaient d'un inventaire à jour des données personnelles stockées et des moyens de stockage employés, que seulement 22% avaient une idée précise des tierces parties qui avaient accès aux données sensibles de l'organisation, et que tout juste 43% procédaient à des évaluations internes de la qualité des processus de gestion des données personnelles (PriceWaterhouseCoopers 2008).

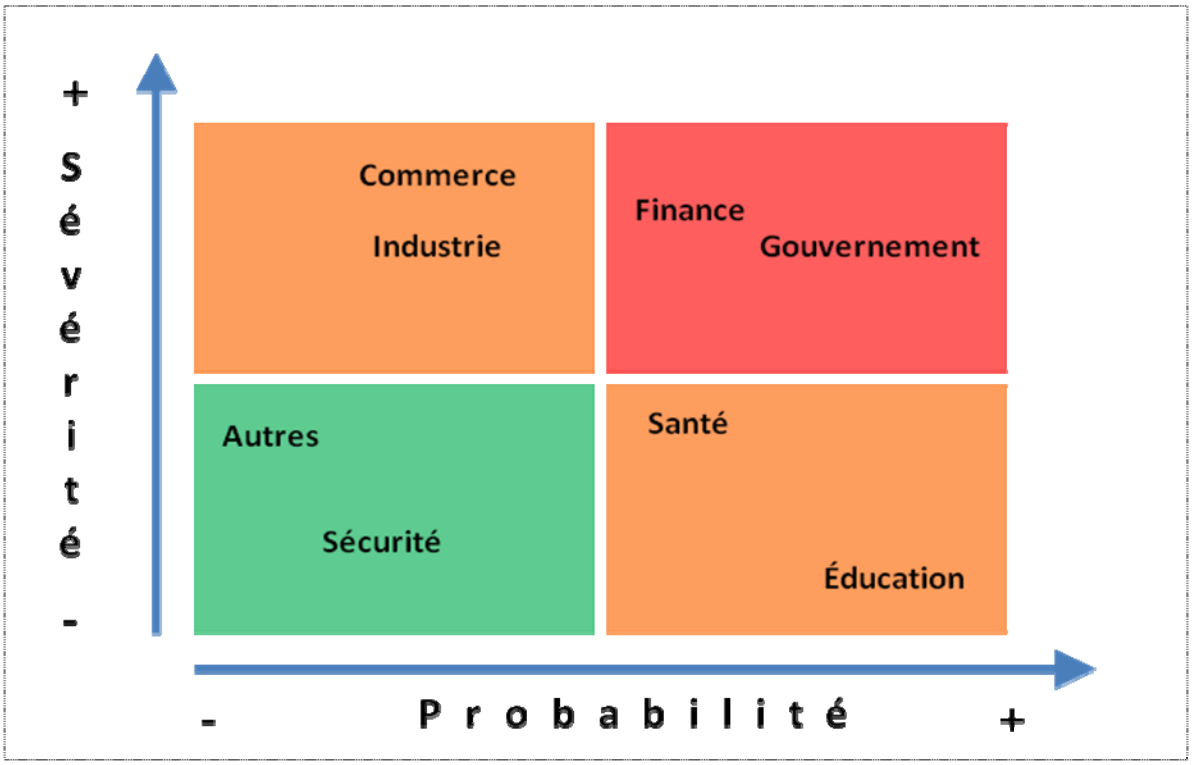

Figure 1 : Quadrant des risques liés à la gestion des données personnelles

b. Diversifier les techniques de prévention

Dans un deuxième temps, cette analyse permettra d'aligner la stratégie de prévention sur le portfolio des risques en faisant appel à l'approche de la prévention situationnelle. Celle-ci s'avère particulièrement attractive dans la mesure où elle permet de concevoir la prévention comme un large éventail d'activités qui s'appuient sur une complémentarité de politiques organisationnelles, de solutions technologiques et de mesures de responsabilisation des individus. Élaborée par le criminologue Ron Clarke et ses collègues, dans une perspective résolument utilitariste (Clarke 1997), cette approche met l'accent sur la structure des motivations et des opportunités qui caractérisent les actes malveillants ou négligents en ne se bornant pas à l'étude du délinquant, mais en y intégrant également les caractéristiques de la cible et la performance des 'gardiens'.

Augmentation des efforts	**Augmentation des risques**	**Réduction des bénéfices attendus**	**Suppression des justifications**
1. Durcissement des cibles	5. Détection des intrusions	9. Camouflage des cibles	13. Activité réglementaire
2. Contrôle d'accès	6. Surveillance technique	10. Perturbation des marchés clandestins	14. Contrôle des «désinhibiteurs»
3. Protection de l'intégrité des données	7. Surveillance par les employés	11. Réduction des tentations	15. Incitations au respect des règles
4. Authentification de l'identité	8. Surveillance par les partenaires	12. Déni des bénéfices	16. Attribution de responsabilité

Tableau 1 : 16 techniques de prévention situationnelle (Newman et Clarke 2003)

Le tableau 1 expose les 16 techniques de prévention qui peuvent être mobilisées pour répondre à ce type de risques, regroupées en 4 grandes catégories : l'augmentation des efforts, l'augmentation des risques, la réduction des bénéfices attendus et la suppression des justifications. Chacune des 16 techniques renvoie à des solutions particulières que nous n'avons pas le temps d'examiner en détail ici, mais qui peuvent prendre un aspect technique (logiciels de cryptage, audit de l'accès aux bases de données), politique (directives d'encadrement du partage des données avec des organisations tierces) ou cibler directement les employés de l'organisation (formation à

la valeur des données personnelles, engagement écrit annuel à respecter les politiques de l'organisation, etc.).

c. Les risques de la gestion des risques

La diversité des mesures de prévention envisageables et le recours simultané à plusieurs d'entre elles exposent toutefois l'organisation au défi de la cohérence. En effet, il n'est pas rare de voir des mesures contradictoires être adoptées en toute bonne foi, pour aboutir à des résultats tout à fait contre-productifs. Ainsi, aux États-Unis, le *Better Business Bureau* recommande-t-il aux organisations de remplacer la facturation papier par la facturation électronique, dans le but de prévenir le vol de courrier dans les boîtes aux lettres. Cependant, on conditionne de la sorte les clients des institutions financières à recevoir des courriers électroniques de la part de leur banque, ce qui les conduit ainsi involontairement à baisser leur garde contre les pratiques d'hameçonnage (*phishing)*.

Dans d'autres cas, c'est l'optimisation qui devra être soigneusement pensée. En effet, la logique de sécurité est rarement celle qui domine dans une organisation, spécialement lorsqu'elle rentre en conflit avec les logiques concurrentes de rentabilité, de compétitivité ou encore de qualité du service offert à la clientèle. Si l'on considère la gestion des accès par exemple, une approche trop restrictive donnera lieu à des violations systématiques de la part de secrétaires ou d'employés qui devront pouvoir accéder aux comptes et aux données réservées des cadres ou de leurs collègues en l'absence de ces derniers. Des mesures trop laxistes affaibliront, quant à elles, la crédibilité des stratégies de protection des renseignements personnels en révélant leur nature essentiellement symbolique. Chaque mesure de prévention doit, par conséquent, être conçue et implantée en tenant compte de ce critère d'équilibre optimal entre efficacité de la sécurité et facilité d'usage dans un contexte de productivité organisationnelle spécifique à chaque milieu.

Conclusion

La sécurité objective des données personnelles ne fait à l'heure actuelle l'objet d'aucune statistique officielle dans les pays développés, en dépit d'indices concordants de la piètre performance des organisations publiques et privées dans ce domaine. À travers une analyse descriptive des données disponibles pour l'Amérique du Nord (la seule région du monde où cet exercice est possible à l'heure actuelle), il a été possible de lever une part du voile sur les risques menaçant l'intégrité des informations détenues par les organisations, qu'il s'agisse de celles concernant leurs clients, leurs usagers ou encore celles de leurs employés. Paradoxalement, c'est aux Etats-Unis, l'un des pays démocratiques où la protection de la vie privée est la moins réglementée que la transparence (plus ou moins forcée) des organisations sur ce sujet est la plus développée. Outre l'atteinte à la réputation et à l'image d'une organisation que représente la divulgation d'un incident relié à la perte ou au vol de données personnelles, les coûts associés à la gestion *ex post* d'une crise dans ce domaine

(communications individuelle avec les victimes, enquête, amendes et pénalités, frais de reconstitution des données, etc.) rendent les investissements en prévention de plus en plus avantageux. Toutefois, la prévention ne peut se limiter au déploiement de technologies miracles conçues sans véritablement tenir compte ni des risques spécifiques auxquels sont exposées les organisations, ni des dynamiques individuelles et collectives qui régissent leur fonctionnement. En outre, ces risques correspondent rarement à l'image véhiculée par les médias et les prescripteurs de solutions de sécurité : s'il existe bien des groupes de pirates informatiques très compétents et déterminés capables de s'emparer de millions de numéros de cartes de crédit, le principal danger vient aujourd'hui plutôt de la négligence et de l'insouciance de ceux qui sont chargés de protéger les données qui leurs sont confiées. À défaut d'une prise de conscience par les organisations de la valeur de ces données dans une société où l'identité se dématérialise un peu plus chaque jour, et sans une amélioration significative de la sécurité que celles-ci sont capables de garantir aux identités numériques, l'opinion publique et les défenseurs de la vie privée pourront légitimement réclamer des lois plus contraignantes et la mise en œuvre de sanctions beaucoup plus strictes.

Références

Mark Button (2008), *Doing security: Critical reflections and an agenda for change*, Palgrave Macmillan, Basingstoke.

Ronald Clarke (ed.) (1997), *Situational crime prevention: Successful case studies*, Harrow and Heston, New York.

Maurice Cusson (2002), *Prévenir la délinquance : Les méthodes efficaces*, Presses Universitaires de France, Paris.

Benoît Dupont et Benoît Gagnon (2008), *La sécurité précaire des données personnelles en Amérique du Nord : Une analyse des statistiques disponibles*, Chaire de recherche du Canada en sécurité, identité et technologie, Montréal.

Graeme Newman et Ronald Clarke (2003), *Superhighway robbery: Preventing e-commerce* crime, Willan, Cullompton.

Ponemon Institute (2008), *Airport insecurity: The case of lost and missing laptops*, Ponemon Institute, Traverse City.

Kieran Poynter (2008), *Review of information security at HM Revenue and Customs final report*, HMSO, Londres.

PriceWaterhouseCoopers (2008), *Safeguarding the new currency of business: Findings from the 2008 global state of information security study*, PWC, Londres.

Rita Tehan (2007). *Data Security Breaches: Context and Incident Summaries - CRS Report RL33199*, Congressional Research Service for Congress, Washington.