

Cahier des charges de l'étude sur le courrier électronique et le « spam »

Ce cahier des charges se compose :

- d'une note précisant les grandes lignes de l'étude ;
- d'un projet de questionnaire pour une enquête par sondage à mener auprès d'un échantillon d'internautes (annexe I) ;
- d'un projet d'exploitation des messages recueillis auprès des internautes de l'échantillon qui utilisent le courrier électronique (annexe II).

Ce projet a été élaboré dans le cadre du groupe de travail des acteurs de la lutte contre le « spam » mis en place par le gouvernement.

Projet d'étude sur le courrier électronique et le « spam »

Ce projet vise à mieux comprendre le « spam », phénomène qui affecte le fonctionnement du courrier électronique par l'expédition de messages que leurs destinataires n'ont généralement pas sollicités et qu'ils ne souhaitent plus recevoir.

Les enquêtes, mesures et analyses existantes manquent en effet d'éléments factuels qui permettraient de mieux cerner toutes les dimensions d'un phénomène complexe, car n'étant pas une réalité homogène. A ce jour, la très grande majorité des données existantes ne portent que sur la proportion de « spams » dans le courrier telle qu'elle est déclarée par l'internaute ou mesurée par des professionnels qui filtrent le « spam », auprès de leurs clients.

Conduites selon des techniques et des méthodes variées, ces opérations fournissent des mesures et des conclusions souvent difficiles à relier entre elles, ce qui ne facilite pas la compréhension du « spam » ni l'appréciation de son impact sur le courrier électronique et l'utilisation de l'internet.

D'où le besoin d'une mesure nouvelle incluant une description factuelle complète du courrier.

Les premiers objectifs du projet sont donc :

- l'établissement d'une description factuelle et détaillée du courrier reçu par les internautes (pour cette description, on ne recourra pas à la notion de courrier sollicité ou non, et donc pas à la notion habituelle de « spam ») ;
- l'observation du comportement sur l'internet et le relevé des opinions sur le « spam » des utilisateurs du courrier électronique. Cette observation comprend des évaluations spontanément déclarées par l'internaute du volume de son courrier et de la proportion de « spams » dans ce courrier qui fourniront une mesure globale de l'intensité de la nuisance ressentie¹ par les internautes.

L'établissement et le rapprochement de ces deux séries d'indicateurs devraient permettre de mieux comprendre ce que les internautes ressentent comme du « spam » (des offres publicitaires, des escroqueries, des courriers au contenu douteux voire illicite, etc.), de mieux apprécier le niveau de nuisance générée par les « spams » et de savoir comment les internautes s'organisent pour gérer ou réduire cette nuisance.

Une attention particulière sera portée à l'analyse des dispersions par rapport aux moyennes et en particulier à celle des déciles extrêmes.

¹ Le projet est centré sur l'internaute et se limite au courrier qu'il a effectivement reçu. Il ne couvre pas les courriers non délivrés à tort comme les « faux positifs » ou à raison comme les courriers non désirés ; cette approche est donc focalisée sur le « spam internaute » et non sur le « spam tuyau ». Ce dernier aspect du sujet demanderait, pour être traité, une approche tout à fait différente de celle proposée ici, qui serait dirigée vers les opérateurs gérant le courrier électronique.

1. Détermination et délimitation de l'ensemble à observer

L'étude concerne les internautes ou plus précisément les usagers de l'internet qui accèdent au réseau par leurs propres moyens (sont donc exclus les salariés n'ayant accès à l'internet que par l'intermédiaire de leur employeur) ; ces internautes sont résidents en France et âgés de 15 ans ou plus.

L'étude concerne plus particulièrement ceux de ces internautes qui sont titulaires d'au moins une adresse de courrier électronique ; pour ces derniers, l'ensemble du courrier qu'ils auront reçu pendant une période de temps déterminée sur l'ensemble des adresses qu'ils utilisent régulièrement, sera analysé selon le protocole d'enquête décrit ci-après².

2. Méthode d'observation

Le projet se compose d'une enquête classique destinée à être administrée à un échantillon d'environ 1.000 internautes et d'une opération de recueil et d'analyse du courrier que ces internautes reçoivent sur leurs adresses personnelles.

Le mode de soumission du questionnaire (téléphone, en ligne, papier, mixte, etc.) sera proposé par le contractant, compte tenu de l'objectif retenu en matière de taux de réponse et de la contrainte financière. Le sujet de l'étude étant un sujet sensible notamment pour une partie de la population des internautes, un soin particulier sera apporté à la maîtrise de la représentativité de l'échantillon des répondants.

Un système de recueil du courrier sera mis en place auprès des internautes participant à l'enquête. Ce système fonctionnera pendant une période liée au calendrier d'exécution de l'enquête d'une durée suffisante pour collecter un échantillon représentatif du courrier reçu par les internautes.

La retransmission du courrier personnel ne sera pas nécessaire et pourra être évitée si l'internaute s'y oppose : en effet, ce courrier que l'internaute ne voudrait pas confier à un tiers, n'est sûrement pas qualifiable de « spam » et est donc hors champ de l'étude. Toutefois, il faudra prévoir un système de comptage de ces messages non transmis afin de pouvoir disposer d'une mesure de l'ensemble du courrier reçu au cours de la période étudiée.

3. Le questionnaire et l'analyse des boîtes aux lettres

3.1. Le projet de questionnaire tel qu'il figure dans l'annexe I de ce document doit être considéré comme l'expression des besoins d'informations du commanditaire de l'étude : basé sur des modèles existants, au niveau européen notamment, pour réserver des possibilités de comparaison internationale, il a été mis au point en laissant ouverts les choix techniques liés au mode d'administration du questionnaire, par exemple. Il reviendra au prestataire retenu de finaliser la formulation du questionnaire et de le tester avant réalisation de l'enquête en vraie grandeur.

² L'enquête se focalise sur l'observation du courrier électronique par rapport à la sphère privée. L'extension de cette observation à la sphère professionnelle aurait en effet poser des problèmes importants d'accès à l'information et aux données, notamment pour les personnes salariées. Par contre, il sera possible d'obtenir des résultats sur la partie du courrier professionnel qui relève de personnes travaillant à leur compte.

Sous sa forme actuelle, le questionnaire comprend quatre modules qui portent sur :

- les caractéristiques générales d'équipement et d'usage au domicile de l'internaute ;
trois questions situent l'équipement du ménage auquel appartient l'internaute, puis une première série de treize questions portent sur la compétence en matière d'utilisation de l'internet, le degré d'usage et le lieu de connexion au réseau (domicile, travail, école), l'objet et la nature des activités en ligne, notamment les commandes de produits et de services passées sur l'internet. Elles devraient permettre notamment de repérer des comportements à risques en matière de communication ou de commerce électronique ;
- le courrier électronique : équipement et utilisation ;
ce module comprend six questions portant sur les adresses utilisées avec leurs principales caractéristiques, l'intensité du courrier reçu quotidiennement avec une ventilation sommaire par grande catégorie et une indication de la proportion de « spams », les temps moyens de traitement du courrier et du « spam ». Les caractéristiques des adresses devraient permettre d'analyser d'éventuels facteurs d'exposition au risque (ancienneté de l'adresse, type de messagerie, suffixe du nom de domaine, protection ou non ...).

La déclaration de l'intensité moyenne habituelle du courrier quotidien doit permettre de s'assurer que la période retenue pour le relevé du courrier n'est pas totalement atypique.

La ventilation du courrier par grande catégorie est demandée indépendamment de la proportion de « spams », les courriers non sollicités pouvant a priori être de toutes natures ;

- le traitement des « spams » ;
ce module comporte quatorze questions qui touchent directement au « spam » sous l'angle de l'identification de ce type de courrier par l'internaute, des conséquences de cette nuisance, de la détermination des caractéristiques du « spam » par l'internaute et de ses éventuelles réactions en réponse à la réception de « spams » ou de courriers non sollicités ;
- les caractéristiques socio-démographiques du répondant (sexe, âge, activité, etc.).

3.2. Le deuxième volet du projet d'étude concerne l'analyse des boîtes aux lettres (BAL) des internautes. Ce volet original de l'étude est particulièrement important à deux titres : d'abord, il doit fournir des éléments factuels sur la composition du courrier électronique qui n'existent pas aujourd'hui ; en outre, il devrait être possible de croiser les données factuelles ainsi obtenues avec les informations apportées par l'enquête sur l'équipement, le comportement et les opinions des internautes.

Ces croisements devraient ouvrir des opportunités d'analyses inédites ; toutefois, elles ne seront pas aisément accessibles car il ne sera certainement pas possible d'exploiter rapidement l'intégralité des milliers, voire dizaines de milliers de messages collectés. D'où, le schéma suivant proposé pour l'exploitation et l'analyse des données relatives aux BAL.

Dans un premier temps, seul un échantillon aléatoire d'un millier de messages sera codifié et exploité afin de fournir les grandes caractéristiques du courrier électronique recueilli en même temps que seront exploités les questionnaires retournés par les internautes. Ces deux exploitations « juxtaposées » des messages et des questionnaires viseront à fournir rapidement les principales données de cadrage.

Ce premier traitement d'un échantillon de messages servira à mettre au point et à valider les propositions de classification présentées dans l'annexe II³ ; il pourra aussi servir à paramétrer un logiciel au moyen duquel sera automatisée l'exploitation de l'ensemble des courriers recueillis auprès des internautes participant à l'enquête.

Cette exploitation plus large fournira alors les éléments nécessaires à l'analyse des liens éventuels entre les caractéristiques et les opinions des internautes d'une part, et celles de leur courrier d'autre part.

4. Le contenu de la prestation et les résultats attendus

Le prestataire réalisera les travaux suivants :

- le recrutement d'un échantillon représentatif d'internautes résidents de 15 ans ou plus ;
- la mise au point finale du questionnaire incluant un test de validation ;
- la mise en place du système de recueil du courrier ;
- la réalisation de l'enquête ;
- les exploitations de l'enquête et d'un échantillon limité du courrier des internautes ;
- le rendu d'un premier rapport présentant les principales tabulations issues des exploitations précédentes ;
- l'exploitation complète des courriers recueillis ;
- le rendu d'un rapport de l'étude comprenant une première analyse du rapprochement entre les questionnaires et le contenu des BAL.

Les données recueillies seront cédées au département statistique de la DDM qui pourra les utiliser pour des analyses supplémentaires.

Trois réunions seront organisées pour présenter à un comité de pilotage, le projet de questionnaire et les outils envisagés pour le traitement des BAL, le rapport intermédiaire et le rapport final.

5. Le choix du prestataire

Les points essentiels pour la réussite du projet sont :

- la disposition d'un échantillon d'internautes dont la représentativité par rapport au sujet étudié devra être démontrée ;
- la fiabilité d'un système de recueil du courrier des internautes ;
- la maîtrise d'un logiciel incluant une fonction d'apprentissage pour extraire les caractéristiques utiles des messages et effectuer les classements recherchés.

La proposition comprendra des indications aussi précises que possible sur ces trois points. Elle donnera une description précise des références en matière d'enquêtes auprès des internautes et d'éventuelles compétences en matière de courrier électronique. Elle sera accompagnée d'une proposition de prix détaillée.

6. Le calendrier

La durée de l'étude sera de trois mois à compter de sa notification.

³ Les éléments nécessaires pour l'exploitation des messages recueillis figurent dans l'annexe II qui présente les principales caractéristiques pouvant présenter de l'intérêt pour l'analyse.

ANNEXE I

Questionnaire pour l'enquête

Le document qui suit est un projet qui décrit très précisément les informations dont la collecte est souhaitée. La mise au point finale de l'agencement du questionnaire, de l'intitulé des questions, de la mise en ordre des réponses proposées, etc. sera assurée par la société qui réalisera l'enquête, en fonction notamment du mode d'administration qui sera retenu.

En particulier, si la passation du questionnaire devait se faire par téléphone, il faudra peut-être aménager la question A8 sur les usages de l'internet au cours des douze derniers mois. De plus, il y aura avantage, pour certaines questions, à proposer les modalités de réponse dans un ordre aléatoire, par exemple pour les questions A8 (mettre l'aléa sur chacun des sous-ensembles), A10 bis, A10 ter, A13, A16, C1, C4, C5, C6 et C6bis, C11 et C12 (par sous-ensemble), C13 et C14. Cette possibilité d'aléa dans l'ordre des réponses proposées est, selon les spécialistes, un avantage qui peut être déterminant pour la qualité (absence de biais) des réponses.

Par ailleurs, des formulations alternatives ont été proposées pour quelques questions. Ces choix restent à effectuer en fonction de considérations techniques : c'est le cas des questions C3, C6 et C6 bis.

A. Caractéristiques générales d'équipement et d'usage au domicile de l'internaute

L'équipement du ménage auquel appartient l'internaute

A1. Sur quels appareils avez-vous accès à l'internet ?

(cochez tous les accès)

- micro-ordinateur
- téléphone mobile ou PDA⁴
- autres (préciser) ...

A2. Type de connexion à l'internet

- haut débit
- bas débit
- je ne sais pas

A3. Type d'abonnement

- au forfait / illimité
- à la durée
- je ne sais pas

Questions concernant l'internaute

A4. Depuis quand utilisez-vous l'internet ?

- cinq ans ou plus
- de deux ans à cinq ans
- moins de deux ans

A5. Sur une échelle allant de 1 à 10, pouvez-vous noter votre niveau de compétence en matière d'utilisation de l'internet ?

A6. En moyenne, quelle est votre fréquence d'utilisation de l'internet (y compris l'accès au courrier électronique) ?

- plusieurs fois par jour
- chaque jour ou presque
- au moins une fois par semaine
- au moins une fois par mois
- moins d'une fois par mois

A7. Où vous connectez-vous à l'internet ? (cochez tous les lieux d'accès)

- au domicile
- au lieu de travail (si distinct du domicile)
- à l'école, au lycée, à l'université
- autres endroits (lieux publics, café internet, etc.)

⁴ Personal Digital Assistant. Les assistants numériques personnels (PDA) sont des systèmes de poche conçus à l'origine comme des organisateurs personnels, mais qui au fil des années sont devenus de plus en plus polyvalents ; le PDA de base comprend habituellement une montre, un agenda, un carnet d'adresses, un aide-mémoire, un bloc-notes et une simple calculatrice.

Objet et nature des activités sur l'internet

A8. Au cours des douze derniers mois, avez-vous utilisé l'internet pour ...
(cochez toutes les activités concernées)

communiquer

- envoyer/recevoir des méls
- téléphoner
- utiliser un logiciel de messagerie instantanée ou participer à des chats
- participer à des forums de discussion
- s'abonner à des listes de mailing, des newsletters

chercher de l'information ou accéder à des services en ligne

- chercher de l'information sur des personnes ou des entreprises
- chercher de l'information sur des produits ou des services
- se renseigner pour organiser un voyage
- écouter la radio / regarder la télévision
- jouer en ligne
- télécharger des fichiers (y.c. logiciels)
- lire la presse en ligne
- chercher un emploi

commander et vendre des produits et des services

- faire des opérations bancaires ou financières
- acheter / commander des produits et des services (non financiers)
- acheter / vendre de particulier à particulier

effectuer des démarches administratives

- obtenir des informations sur des sites administratifs
- télécharger des formulaires officiels
- renseigner un formulaire en ligne

se former et s'éduquer

- dans le cadre de la formation initiale (école, université ...)
- dans le cadre de la formation continue

la santé

- rechercher de l'information médicale

A9. Au cours d'une des utilisations précédentes, avez-vous déjà été amené à communiquer des données personnelles vous concernant sur un site ?

oui non je ne sais pas

(Si NON, passer directement à la question A12.)

A10. Si OUI, combien de fois environ au cours des douze derniers mois ? —

A10 bis. De quelles données s'agit-il ?

(cochez toutes les réponses vous concernant)

mon adresse de courrier électronique —

mes coordonnées bancaires —

mon numéro de carte bancaire —

mes coordonnées postales —

autres ... (préciser) —

A10 ter. Dans quel contexte vous a-t-on demandé de fournir ces données ?

(cochez toutes les réponses vous concernant)

pour accéder à des informations sur un site —

pour vous abonner à une lettre d'information —

pour effectuer un paiement en ligne —

pour obtenir une livraison à domicile —

autres ... (préciser) —

A11. A combien de lettres d'information (newsletters) êtes-vous abonné ?

aucune —

une à cinq —

plus de cinq —

A11 bis. Si vous êtes abonné à des lettres d'information,
certaines de ces lettres sont-elles payantes ?

oui _ non _

A12. De quand datent vos derniers achats pour votre usage personnel sur l'internet ?

de moins de trois mois —

de plus de trois mois à moins d'un an —

de plus d'un an —

n'a jamais acheté ni commandé —

(Si vos achats datent de plus d'un an, passer directement à la partie B ; si vous n'avez jamais acheté ni commandé, passer directement à la question A15.)

A13. Depuis douze mois, quels produits et quels services avez-vous commandé sur l'internet pour votre usage personnel ? (cochez tous les produits et services concernés)

nourriture / épicerie —

films, musique —

livres / magazines / journaux / matériel éducatif —

vêtements / articles de sport —

produits d'hygiène / de santé —

logiciels informatiques et jeux vidéo —

matériel informatique —

matériel TV, Hi-fi, photo —

ordres de bourse / services financiers / assurance —

billet de voyage ou location de vacances —

billet pour un spectacle —

billet de loterie ou pari —

biens domestiques (meubles, jouets ...)

autres ... (préciser) —

A14. Avez-vous commandé ces produits ou services ... ?

(cochez toutes les réponses vous concernant)

- sur des sites francophones
- sur des sites anglophones
- sur d'autres sites*
- je ne sais pas

(*) c'est-à-dire des sites édités dans une langue autre que le français ou l'anglais.

A15. Si vous n'avez jamais acheté ni commandé de produits ou de services sur l'internet, avez-vous l'intention d'acheter ou de commander dans un proche avenir ?

oui non je ne sais pas

A16. Et si vous n'avez pas l'intention d'acheter ou de commander sur l'internet, c'est pour quelles raisons ? (cochez toutes les raisons concernées)

- ça ne m'intéresse pas
- les produits ou services proposés ne m'intéressent pas
- je crains des livraisons non conformes à la commande
- je ne veux pas communiquer mes coordonnées personnelles
- je ne veux pas donner mon numéro de carte de crédit
- autres ... (préciser)

B. Le courrier électronique : équipement et utilisation

Le tableau suivant porte sur les adresses qui vous permettent de recevoir « habituellement » votre courrier électronique (*il est précisé « habituellement » car on ne prendra pas en compte les adresses qui ne seraient plus utilisées⁵.*)

Remplir le tableau suivant, S.V.P.

B1. Vos adresses personnelles

N°	Adresses	Ancienneté	Mode d'accès	Fonction	Protection
1					
2					
3					
4					
5					

Pour la description de chaque adresse, préciser :

- l'ancienneté (en nombre d'années ou en nombre de mois, si l'adresse date de moins d'un an) ;
- le mode d'accès, par un « client de messagerie » (logiciel de type Outlook ou Eudora ...) ou par un « navigateur » (logiciel Internet Explorer) ;
- la fonction (indiquer au moins la distinction courrier personnel / courrier professionnel / mixte, personnel et professionnel s'il y a lieu, et ajouter d'éventuelles spécialisations prévues par l'internaute, par exemple, adresse réservée au courrier publicitaire, etc.) ;
- l'existence ou non d'une protection « anti-spam ».

Pour ces adresses, la société qui réalisera l'enquête récupérera le courrier reçu par l'internaute hors le courrier personnel dont il ne souhaite pas se dessaisir. Des compteurs seront installés pour faire le comptage quotidien de tous les courriers reçus (avec, si c'est possible, une indication du volume en Ko).

B2. Au total, combien recevez-vous de méls par jour en moyenne, sur chacune de vos adresses⁶ ? (cochez une case pour chaque adresse)

N°	moins de 5	6 à 10	11 à 15	16 à 20	21 à 30	31 à 50	51 à 100	plus de 100	je ne sais pas
1									
2									
3									
4									
5									

⁵ La distinction sera faite par l'internaute.

⁶ Pour les adresses filtrées, le volume de courrier sera considéré avant filtrage.

B3. Pouvez-vous ventiler ce courrier selon les grandes catégories suivantes ?

N°	Messages personnels	Messages professionnels	Messages publicitaires	Autres	Total
1					100%
2					100%
3					100%
4					100%
5					100%

B4. Dans tout ce courrier, à combien estimez-vous la proportion de « spams » c'est-à-dire de courriers que vous n'avez pas sollicités et que vous ne voudriez plus jamais recevoir ?
(cochez une case pour chaque adresse)

N°	nulle	moins de 10%	de 10% à moins de 25%	de 25% à moins de 50%	de 50% à moins de 75%	75% ou plus	je ne sais pas
1							
2							
3							
4							
5							

B5. Au total, combien de temps estimez-vous passer en moyenne par jour à traiter⁷ votre courrier électronique ?

- aucun —
- moins de 5 minutes —
- de 5 à 15 minutes —
- de 15 minutes à moins d'une demi-heure —
- d'une demi-heure à moins d'une heure —
- une heure ou plus —
- je ne sais pas —

B6. ... et combien à traiter les « spams » ?

- aucun —
- moins de 5 minutes —
- de 5 à 15 minutes —
- de 15 minutes à moins d'une demi-heure —
- d'une demi-heure à moins d'une heure —
- une heure ou plus —
- je ne sais pas —

⁷ Par « traiter », on entend « recevoir, prendre connaissance et répondre ».

C. Le traitement des « spams »

C1. Pour éviter de recevoir des « spams », avez-vous déjà eu recours aux procédés suivants ? *(cochez tous les procédés utilisés)*

- éviter de diffuser votre adresse mél
- éviter de faire apparaître votre adresse mél sur un site web
- créer une adresse mél inhabituelle difficile à deviner
- utiliser un filtre « anti-spam »
- utiliser une adresse particulière chaque fois qu'il y a un risque
- donner une fausse adresse
- autres ... *(à préciser)* ...
- je ne fais rien

C2. Globalement, considérez-vous les procédés précédents comme efficaces ? *(une seule réponse)*

- oui en majorité, ils sont efficaces
- ils sont efficaces en partie seulement
- non, ça ne marche pas
- je ne sais pas

C3. Laquelle de ces affirmations décrit le mieux comment vous identifiez un « spam » ? *(une seule réponse)⁸*

- je le reconnais sans l'ouvrir
- j'ai souvent du mal à le distinguer d'un autre mél
- je ne sais pas

C4. Comment identifiez-vous le « spam » dans votre courrier ? *(cochez tous les procédés utilisés)*

- d'après la langue utilisée
- d'après l'expéditeur
- d'après « l'objet »
- d'après la présentation du message (texte, html ...)
- d'après le contenu
- autre procédé ... *(à préciser)* ...
- je ne sais pas

C5. Habituellement que faites-vous d'un courrier que vous pensez être un « spam » ? *(cochez tous les procédés utilisés)*

- je le détruis immédiatement sans l'ouvrir
- je l'ouvre et le vois rapidement avant de le détruire
- je le signale à mon fournisseur d'accès
- je clique pour être rayé de la liste d'envoi
- je clique pour avoir plus d'information
- autre chose ... *(à préciser)* ...

⁸ Formulation alternative possible de C3 :

Arrivez-vous à identifier facilement le « spam » dans votre boîte aux lettres ?

oui non parfois

Le choix final sera fait avec la société qui réalisera l'enquête en fonction de l'efficacité de la formulation.

C6. Les « spams » ont-ils une influence sur votre utilisation de l'internet ?
(donnez une réponse par ligne)⁹

	OUI	NON	je ne sais pas
en réduisant votre utilisation du courrier électronique	—	—	—
en augmentant le temps passé à lire et traiter les e-mails	—	—	—
en diminuant votre confiance dans le courrier électronique	—	—	—
en compliquant l'utilisation de l'internet	—	—	—

C6 bis. Si OUI, l'effet des « spams » sur votre utilisation de l'internet est-il important ou limité ? (donnez une réponse par ligne)

	effet : important	limité
en réduisant votre utilisation du courrier électronique	—	—
en augmentant le temps passé à lire et traiter les emails	—	—
en diminuant votre confiance dans le courrier électronique	—	—
en compliquant l'utilisation de l'internet	—	—

C7. Craignez-vous de perdre un courrier qui vous est adressé à cause d'un filtre « anti-spam » ?

oui _ non _ je ne sais pas _

C7 bis. Cela vous est-il déjà arrivé ?

oui _ non _ je ne sais pas _

C8. Craignez-vous qu'un de vos courriers n'arrive pas à son destinataire à cause d'un filtre « anti-spam » ?

oui _ non _ je ne sais pas _

C8 bis. Cela vous est-il déjà arrivé ?

oui _ non _ je ne sais pas _

C9. Dans un futur proche, avez-vous l'intention de vous équiper d'un système de filtrage « anti-spam » ?

oui _ non _ je ne sais pas _

C10. Si aujourd'hui, vous souscriviez une nouvelle adresse de courrier électronique, prendriez-vous en compte la politique « anti-spam » du FAI envisagé ? (une seule réponse)

oui, à mes yeux c'est un critère important —

c'est l'un des critères que je vais prendre en compte parmi d'autres —

non, ça ne me paraît pas un critère majeur —

je ne sais pas —

⁹ Cette question et la question qui suit peuvent être fondues en une seule, les modalités des réponses devenant alors « OUI, effet important, effet limité » au lieu de « OUI, NON, je ne sais pas ». C'est un point de présentation pour lequel sera pris l'avis de la société qui réalisera l'enquête.

C11. Considéreriez-vous les courriers non sollicités suivants comme du « spam », par exemple ... (cochez les « spams »)

des e-mails commerciaux expédiés par quelqu'un ...			
à qui je n'ai sûrement jamais communiqué mon adresse			—
que je ne peux identifier			—
que je peux identifier			—
à qui j'ai récemment acheté quelque chose			—
ou des e-mails provenant ...			
d'un organisme associatif ou caritatif			—
d'un parti politique ou d'un syndicat			—
d'un expéditeur ou d'une société autorisée à vous contacter			—
de quelqu'un que vous ne connaissez pas			—
et qui s'adresse personnellement à vous			—

C12. Considéreriez-vous les courriers non sollicités suivants comme du « spam », par exemple ... (qualifiez chaque exemple d'e-mail)

	OUI, toujours	quelques fois	NON, jamais
des e-mails qui proposent ...			
un placement financier	—	—	—
du contenu pour adultes	—	—	—
un logiciel	—	—	—
un produit ou un service	—	—	—
ou des e-mails qui contiennent ...			
de l'information religieuse	—	—	—
de l'information médicale	—	—	—
de l'information politique ou syndicale	—	—	—
de bonnes histoires, des plaisanteries,			
de quelqu'un que vous connaissez	—	—	—
des fichiers, de quelqu'un que vous connaissez	—	—	—

C13. En réponse à un courrier que vous avez considéré comme un « spam », avez-vous déjà ... ?¹⁰ (cochez toutes les propositions pertinentes)

commandé un produit ou demandé un service	—
cliqué sur un lien pour en savoir plus	—
communiqué votre adresse électronique	—
donné des informations personnelles autres que l'adresse	—
fait un don ou effectué un versement	—
demandé à être rayé d'une liste d'envoi	—
porté plainte auprès de votre fournisseur d'accès	—
signalé ce « spam » auprès d'organismes qui les répertorient ¹¹	—
contacté une administration (CNIL, DGCCRF, ART, police ...)	—
envisagé sérieusement de porter l'affaire en justice	—

¹⁰ La société qui réalisera l'enquête sera chargée de proposer un autre classement des réponses.

¹¹ comme « thisisspam » de Yahoo, « reportSpam » d'AOL, SpamNet, Hotmail ...

C14. Selon vous, le « spam » pour le courrier électronique, c'est ... ?
(cochez toutes les propositions pertinentes)

... la rançon du succès, et il faut faire avec. —

... un usage abusif et il faut l'éliminer. —

... illégal et les « spammeurs » doivent être poursuivis. —

... un phénomène de mode et il s'éteindra de lui-même. —

D. Caractéristiques socio-démographiques générales

D1. Age

D2. Sexe

D3. Niveau d'éducation

Primaire

Secondaire

Supérieur

D4. Situation d'activité

Salarié

A votre compte (y compris aides familiaux)

Chômage

Étudiant

Sans activité (retraité, inactif, etc.)

D5. Profession

D6. Taille du ménage

dont nombre d'enfants de moins de 16 ans

L'exploitation des messages¹²

Rappel du protocole d'enquête. L'organisme enquêteur recueille une copie des messages reçus par les internautes de l'échantillon pendant la durée de l'enquête. La collecte du courrier personnel n'étant d'aucune utilité pour la présente étude, l'internaute pourra la refuser¹³. Toutefois, il faudra prévoir un système de comptage de ces messages non transmis afin de disposer d'une mesure de l'ensemble du courrier reçu au cours de la période étudiée.

Devraient ainsi être recueillis des messages à caractère publicitaire, des lettres périodiques et des courriers au contenu douteux (partie du « spam »), principalement ; mais ce recueil ramènera aussi d'autres types de message y compris des messages personnels qui sont inutiles pour la suite de l'étude (une typologie expérimentale est proposée ci-après).

Compte tenu du volume attendu de messages à traiter (probablement quelques dizaines de milliers), les classements suivants, par format, caractéristique et contenu ne sont envisageables que s'il est possible de les automatiser de manière satisfaisante.

Les formats de messages

(classification sans référence au contenu)

Les messages publicitaires (méls marketing, catalogues, etc.) ;
Les lettres périodiques d'information (newsletters) ;
Les messages adressés automatiquement (accusé de réception, annonce d'une parution, etc.) ;
Les messages incluant un virus ou résultant d'une attaque virale ;
Les autres (à compléter ... par exemple, les factures, etc.).

Les caractéristiques du message

Taille de l'envoi auquel appartient le message (la détermination de ce paramètre dépendra de la capacité à définir et repérer des messages identiques, ayant un même expéditeur ou un même signataire, une même date d'envoi, un même objet, un même contenu ?)

Adresse IP de l'expéditeur (distinguer les routeurs)

Date de réception

Objet du message (notamment à blanc, falsifié ...)

Langue du message (tenir compte de l'alphabet)

Format informatique du message (texte, html ... contenant une pièce jointe ...)

Taille physique du message (en Ko)

¹² Ce volet de l'étude n'ayant jamais été mis en œuvre est expérimental. Les sociétés candidates établiront une proposition susceptible de répondre aux spécifications présentées dans cette annexe.

¹³ Le « pair à pair » relève du courrier personnel.

Le contenu du message

Coordonnées physiques de l'expéditeur signataire (notamment nom, adresse)

Existence d'une option de désabonnement

Existence d'une adresse de retour

Description du contenu (en particulier, description de l'action que le spam cherche à déclencher, par exemple site vers lequel le courrier propose d'emmener le destinataire ...)

Un élément important de cette exploitation serait d'arriver à distinguer clairement les messages licites et illicites.

Pour une première expérimentation, il est proposé de tester la nomenclature suivante basée sur les caractéristiques des messages filtrés ou recueillis aux États-Unis ; elle devra bien entendu être adaptée par la suite à la réalité de ce qui sera recueilli en France.

Nomenclature de contenu

(à mettre au point avec les messages collectés)

PUBLICITÉ

INFORMATIQUE/INTERNET

VOYAGES/TOURISME

SERVICES FINANCIERS

JEUX

ENSEIGNEMENT/FORMATION

SANTÉ

ADULTES

AUTRES PRODUITS (publicité licite)

PUBLICITE ILLICITE & ESCROQUERIES (« Scams »)

SERVICES FINANCIERS

SANTÉ

ADULTES

AUTRES

AUTRES

Une mesure factuelle du « spam » ? (à tester)

Des solutions au problème du spam sont déjà proposées par des prestataires, fournisseurs d'accès et éditeurs de logiciels, où l'internaute est invité à pointer et à marquer les messages qu'il considère comme étant du « spam » au moyen d'un logiciel de filtrage.

Si ce dispositif fonctionne chez l'internaute, le contractant récupérera le résultat du tri effectué par l'internaute et/ou son système de filtrage. Le but est de savoir s'il est possible d'utiliser cette information résultant du marquage et du tri des messages, pour produire un indicateur du « spam » plus objectif que les indicateurs déjà obtenus par déclaration en faisant appel à l'appréciation et à la mémoire de l'internaute¹⁴.

¹⁴ Dans le projet actuel de questionnaire, cf. la question B4, « Dans tout ce courrier, à combien est-imez-vous la proportion de « spams » c'est-à-dire de courriers que vous n'avez pas sollicités et que vous ne voudriez plus jamais recevoir (comme des messages au contenu douteux ou d'origine virale) ? »