Notes de lecture

iGor

2015-10-22

Contents

1 Définitions

À définir :

- *privacy* et proposer une/des traductions.
- Internet.
- Web.
- *Standard-setting organisations*, standards.
- *reviewing for*.
- IEEE : Institute of Electrical and Electronics Engineers
- CS : Computer Society (association prof depuis 1971, conf, ateliers, peer reviews journals

Nick Doty :

- npdoty.name
- CV
- bachelor en philo.
- master en information management system
- phD en IS
- privacy analyst for W3C (Privacy Interest Group (PING)), Do Not Track

2 Abstract

Fonctionnalité Internet et Web déterminées par Standards → implémentations interopérables. → l'intimité des relations en ligne dépendent du travail fait dans les organistions qui établissent les standards.

Quelle importance des structures et modes de fonctionnement de ces orgs sur des valeurs comme la *privacy* ? [Question de recherche ?]

L'article évalue :

- l'histoire des considérations à propos de la *privacy* et de la sécurité dans ces orgs,
- l'impact de *Snowden* et les réactions,
- tendances sur la manière d'évaluer ces questions (*privacy* et sécurité).

3 Introduction

Fonctionnalités ← standards (permettent des implémentations interopérables). *online privacy* dépend du travail des orgs développant ces standards.

L'évaluation des points liés à la *privacy* sont toujours faits sur les systèmes logiciels, dans le contexte des standards Internet, c'est pas si simple.

Les orgs travaillent par concensus, et la participation est sur une base volontaire, et sur un mode *bottom-up*, contairement au *top-down* des grandes boîtes.

Structure des standards Internet : par couche et *generative* (comme dans grammaire générative). But : permettre une grande variété d'implélmentation, d'application. → résistent à l'analyse systématique (**à préciser**).

Mentionne des travaux précédents présentant ces orgs (*multistakeholder groups*) comme des organisations frontières, avec des réponses innovatives. Leurs structure et approches ne sont pas suffisamment comprises.

Description du plan du travail :

1. II : méthode, data, champ du travail
2. III :
    1. procédures, outils, structure organisationnelle de IETF impactent les aspects de la sécurité, et le rapport avec la *privacy* des standards Internet
    2. contexte des standards spécifiques à la *privacy* pour le Web, l'auteur montre comment ce point est conçu (perçu, pensé ?) et comment la revue (évaluation) de ce point est réalisée
3. IV : description des différentes réactions dans la communauté qui établit les standards face aux révélations Snowden, et l'impacte de celle-ci sur la manière d'évaluer les standards en lien avec la *privacy*.
4. V : Sur base de III et IV, identification de 3 tendances : systématisation, intégration, *leadership*

4 Méthode

4.1 Sources de données

3 sources de données :

1. Les standards Internet et Web eux-mêmes, corpus pour analyse textuelle automatique, indique et confirme les tendances quantitativement. Et l'activité y relative sur les mailings lists publiques.
2. Les articles d'information *mainstream*, rapports de rencontres, principaux documents de standards → constituer la *timeline* et le type de réponse aux révélations Snowden.
3. Entretiens semi-structurés avec des ingénieurs experts d'Internet et qui participent à IETF → perpective interne.

4.2 Champ de recherche

Ne suppose pas une def. de *privacy*, mais : protection contre l'intrusion, immixion (cf. déclaration des droits de l'homme, art. 12) **ET** contrôle des informations concernant soi-même.

essentially-contested concept wp et wpfr : n'a pas donné plus d'importance aux interviewés, ni supposé qu'ils aient une def partagée de la *privacy*. Pourtant, ça a une influence sur leur travail de conception de standards.

Centré sur deux orgs définissant des standards :

- Internet Engineering Task Force(IETF)
- World Wide Web Consortium (W3C)

Les deux incontournables pour les protocols Internet et Web et sur model volontaire et concensus → pas de limite précise sur où se définissent les standards. Ex HTML5 et le WHATWG (Web Hypertext Application Technology Working Group), groupe de vendeurs de navigateurs. En collaboration ou non avec le W3C.

Autre exemple : OASIS, pour des applications du XML en sécurité, santé, web, commerce, etc. (Mais aussi le ODT). Se sont préoccupés de standards pour des méthodologie d'organisation devant gérer des éléments de *privacy* et se penchent sur des standards de processus qui prévoient dans leur conception la *privacy*.

Trop de lieux/orgs qui s'occupent de standards pour tous les mentionner, décrire. Sélectionne IETF et W3C, parce qu'il y accède... cf section *Future Work* qui devrait vérifier si dans les autres orgs définissant des standrards ça fonctionne de manière similaire ou pas.

5 Histoire des évaluations de la *privacy* et de la sécurité

S'intéresse d'abord à la *privacy*, puis aux notions de sécurité en lien à IETF, puis W3C.

5.1 Internet Engineering Task Force

Dispose de l'historique complet des documents publiés depuis 1969. Les publications de standards sont les RFC : *Requests for Comment*. La *privacy* != un sujet explicite, mais présent depuis le début sous le thème de la sécurité des communications.

La sécurité comme précurseur utile de la notion de *privacy* parce que :

- La sécurité, comme *privacy*, a11y, i18n, performance, etc, sont des sujets transversaux ou *horizontaux*.
- Plusieurs aspects de la *privacy*, la sécurité sont des pré-requis (confidentialité, intégrité);
- Tech. en couches, la sécurité des couches de base déterminent si la *privacy* peut être possible à une couche applicative supérieure.
- Non seulement il y a rel. entre sécurité et *privacy*, mais sécurité est un domaine qui peut s'appuyer sur expérience et méthodolgie, qu'on peut adapter pour défendre la *privacy*.

5.1.1 Historique des considérations relatives à la sécurité

Vers 1987, intro d'une obligation d'aborder sécurité dans RFC → augmentation importante, puis 100% des RFCs mentionnent au moins le sujet (voir fig1. Pas le cas pour les mentions de *privacy*, au max 20% environ. L'obligation a un effet...

RFC 1543, 1993, *Instructions to RFC authors*. Toujours mise à jour (encore en 2014). Ne s'occupe que de la mise en forme (modèle), n'impose pas d'élément de contenu en matière de sécurité. Tout le monde d'accord pour dire que durant les 90, les considérations de sécurité étaient insuffisantes dans les RFCs.

Mentionne RFC 3552, 2003, *Guidelines for Writing RFC Text on Security Considerations* :

> All RFCs are required to hava a Security Considerations section. Historically, such sections have been relatively weak.

En 1996, une RFCs signale que les sections sur la sécurité ne sont pas si nombreuses et surtout très brèves...

L'analyse textuelle automatique [on aurait aimé une courte description] confirme, voir fig2. La légende de la figure explique un peu le type d'analyse : *parse* le *plain-text*, repère les sections, et le nombre de ligne à largeur fixe, puis calcul d'un rapport entre la longueur de la RFC et la longueur de la partie "sécurité".

Considérations sécurité quasi absentes avant 90, très brèves au cours des 90. Depuis 2000, les sections se ralongent quelque peu, une plus grande proportion de chaque RFC, mais la plupart des sections restent minimalistes. La longueur des sections sécurité ne garantissent pas qualité sécurité, mais bonne mesure de l'attention portée au sujet.

Éléments technique renforçant la conformité avec exigences pour rédaction RFCs : modèle comprenant une section *Security Considerations* + outils de vérification de soumissions vérifiant la présence des sections. Ces sections sont ensuite évaluées par des reviewers volontaires qui participent au *Security Directorate*.

Selon affirmation d'un interivewé de IETF,

> Now everyone [thinks about security]. Not everyone does, but as soon as you don't, you get called out."

Info donnée aux *Area Directors*, part. les deux *Security Area Directors*. Chaque RFC évaluée par *IESG* (*Internet Engineering Steering Group*), RFCs peuvent être rejetées, ou sujettes à révisions si considérations sur la sécurité ou *privacy* pas suffisantes. Parfois évaluation pointilleuses et approbation difficile

5.1.2 De la sécurité à la *privacy*

RFC 6973, juillet 2013, introduit des recommandations pour la rédactions RFCs en rapport avec *privacy* :

- liste des menaces particulières
- des réduction de ces risques
- une liste de question à vérifier pour identifier et résoudre les problèmes de *privacy*

Existe aussi des RFCs spécifiques à la *privacy* : le `geopriv` Working Group, développement de standards sur transmission des données de géolocalisation (qui prendrait en compte le choix de l'utilisateur, donc pas seulement sécurité des communications). Pendant plusieurs années, le groupe a développé :

- des conditions nécessaires
- analyses de menaces (modèles de menace ?)
- formats de fichiers
- plan

d'un modèle basique de communication d'information de géolocalisation sous forme de standard, avec l'exigence d'une politique claire de l'utilisation de ces données, communiquée à et contrôlée par l'utilisateur. Modèle jugé trop exigeant, existence d'une controverse.

Le fait que ces préoccupations s'appliquent à des domaines de types applicatifs est en cohérence avec les standards attentifs à la *privacy* au W3C : web est une couche applicative d'Internet.

5.2 La *privacy* dans les standards du *World Wide Web Consortium*

IETF : RFCs. Pour le W3C, ce sont des TRs pour *Technical Reports*.

W3C publie aussi standards techniques, en accès libre → analysables. Pas d'exigences spécifiques comme au IETF, mais un nombre plus stable, et relativement important (envir. 20 %), voir fig3. Le graph montre le pourcentage de TR qui mentionne au moins une fois le terme recherché.

5.2.1 Standards spécifiques sur la *privacy*

Au W3C efforts significatifs spécifiques avec *privacy*. Non seulement mentionné ou abordé, mais même des standards qui s'occupent principalement de cette question sur le Web.

5.2.1.1 *Platform for privacy Preferences Project* (*P3P*)

http://www.w3.org/P3P/

Était effort sur plusieurs années d'améliorer la prise de conscience des pratiques des sites Web concernant *privacy* au moyen de description *machine readable* de leurs politiques. Existent d'autres tentatives de résoudre le problème : les divers projets d'icônes pour la *privacy* (http://opennotice.org - pas accessible

le 2015-11-04); efforts avec plusieurs participants, soutenu par le US.gov pour établir des standards pour des avertissements courts, clairs et transparents.

P3P définit un langage XML extensible (?) pour communiquer les pratiques *privacy*, en développant un concept utilisé pour afficher quels contenus sont appropriés pour tel ou tel groupe d'âge. Conçu par couche pour offrir un langage neutre et descriptif pour les pratiques offrant une flexibilité pour les implémentations ou permettant aux usagers de se faire leur propre idée sur les différentes pratiques et sur les différentes possiblités de gérer les préférences en matière de *privacy*. Selon le principe : "*mechanism, not policy*". Pas eu un grand succès (sites Web et navigateurs) : pour cause, du moins exprimée, de trop grande complexité et manque d'incitation.

5.2.1.2 *Do Not Track* (*DNT*)

http://www.w3.org/TR/tracking-dnt/

Effort sur plusieurs années → mécanisme offrant à l'utilisateur le choix face au *tracking* du comportement en ligne. 2010 : pris en charge par *Federal Trade Commission*. 2011 implémentés dans les navigateurs, et efforts de standardisation. L'utilisateur doit pouvoir définir dans les préférences de son navigateur s'il accorde ou non aux sites Web le *tracking*, préférence qui doit être communiqué aux services en ligne. Comme P3P, n'impose pas des propriétés type *privacy*, ne limite pas automatiquement le *tracking*. Suppose coopération entre les *users agents* et les serveurs (qui décideraient de tenir compte des préférences des utilisateurs).

À été mis à jour, plus grande souplesse offerte aux serveurs d'indiquer comment ils tiennent compte des préférences des utilisateurs. Mention du *Privacy Badger* de l'EFF, qui bloque les cookies tiers s'ils n'affichent pas leur méthode DNT.

Autre initiative : DAA (Digital Advertising Alliance) propose sa propre méthode DNT. Mention du group de travail TPWG (*Tracking Protection Working Group*, auquel participe l'auteur) : définit des pratiques en matière de *privacy* auxquelles les sites pourraient adhérer. → diversité des pratiques (politiques) caractéristique d'un manque de standardisation, ou standardisation si étroites (permissives) que possibilité d'une grande diversité d'implémentations.

5.2.1.3 Évaluation de la *privacy* au W3C

Privacy apparaît dans les spécifications de plusieurs APIs et protocols du Web qui ne concernent pas directement la *privacy*. Pas d'exigence de sections "sécurité" ou *privacy* dans les documents publiés (TRs, contrairement à RFCs), mais dans les procédures d'évaluations. Autant à IEFT que W3C, une phase "*Last Call*" ou "*wide review*" : une évalution par un public plus large, en dehors des auteurs et du groupe de travail développant la spécification. Déterminé dans la charte de chaque WG : quel autre WG devra faire des commentaires d'évaluation, souvent sur *privacy*, sécurité, a11y, i18n. Le directeur du W3C peut décider si telle spécification suit le cours normal d'évaluation ou doit être spécifiquement être analysé sur des points particuliers.

Le *PING*, de manière informelle ou de type consultation, apporte son expertise, conseils sur questions de *privacy*, typiquement sur demande d'un WG. PING composé de volontaires : académiques, société civile, industrie, intéressés dans la *privacy* sur le Web. Travaillent également sur docs et procédures pour améliorer l'évaluation de la *privacy*. D'autres groupes W3C apportent leurs commentaires sur *privacy* et sécurité : TAG (*Technical Architecture Group*), *Web Application Security Working Group* et *Web Security Interest Group*.

Privacy pas le seul sujet transversal qui a conduit a évaluer le processus de formation des standards au W3C, parfois mieux formalisés. *Internationalization Working Group* apporte conseil et évalutation dans le cadre W3C et en dehors. Idem pour la *Web Accessibility Initiative* : développe ses propres standards et offre évaluation à d'autres travaux au sein du W3C.

6 Réaction à Snowden

Les révélations Snowden d'une surveillance gouvernementale à grande échelle des communications électroniques à profondément affecté la communauté des standards d'Internet. Les réactions ont été composées de :

- déclarations individuelles / organisations ;
- formations de nouveaux groupes et collaborations ;
- réponses directes sous forme de standards et code.

Grande liste réaction, suffisamment nbreuses pour remplir plusieurs papiers, ici seulement un bref résumé centré sur ce qui impacte l'évaluation de la *privacy* et de la sécurité. Ces évaluations existaient avant Snowden. Mais événement extérieur a inspiré des réponses concrètes et a modifié pratiques des organisation élaborant standards.

1ˆer articles de Greenwald en juin 2013 (en même temps que *Privacy Law Scholars Conference in Berkley*) avec détails sur la section 215, collection des métadonnées téléphoniques et sur l'accès `Prism` aux serveurs des grandes compagnies techniques. Mais encore plus pertinent pour ceux qui travaillent sur la sécurité des connections Internet elles-mêmes, le programme `XKeyscore`, révélé en juillet, et le programme `Bullrun`, révélé en septembre, qui offrent des preuves étonnantes sur les capacités et les pratiques de la NSA en matière de surveillance de l'activité Internet, y compris le trafic chiffré. Encore plus spécific à la communauté élaborant les standards, le même article de sept. dévoile que la NSA a inséré des vulnérabilités dans la sécurité dans le développement d'un standard de chiffrement au *National Institute of Standards and Technology* (NIST).

Des réactions émotionnelles. Documents de 7 pages, texte complet du *A Simple Statement*, par une personne importante au sein du IETF :

> we had a good thing
> you messed it up
> for everyone
> we trusted you

> we were naive
> never again

IETF, nov. 2013, une rencontre pour des déclarations plus larges et de toute l'organisation. Session plénière, plusieurs centaines de participants, demande de supporter ou opposition à la déclaration suivante :

> Pervasive surveillance is an attack, and the IETF needs to adjust our threat model to consider it when developing standards track specifications.

Un fort "hum" de soutien, et un silence en matière d'opposition → consensus confirmé dans un document plus rigoureux détaillant la nature des attaques et les procédures d'atténuation des risques proposées par IETF. En particulier RFC 7258 précise que des considérations au sujet de la menace d'une surveillance intrusive (omniprésente) doit être présente dans les protocols exitants et nouveaux, mais pas une section séparée.

Des individus ont également formé des groupes en réponses. `XKeyscore` a été publié en même temps que le IETF *meeting* à Berlin. Un groupe informel s'est rencontré là, a démarré la *mailing list* `perpass`, forme la plus basique d'orga. à IETF, et une rencontre *BoF* (*birds of a feather*, qui a pour but de mettre en place une activité plus formelle), et un atelier (par IAB – *Internet Architecture Board* – et W3C) pour travailler sur le renforcement d'Internet contre la surveillance intrusive.

fig4 – Stats d'une *mailing list* : aperçu *grossier* sur le niveau d'activité d'une communauté. La liste `perpass` (orange) montre une grande activité en fin 2013, en comparaison avec les autres listes. Montre que cette conversation auto-organisée démarre après les révélations Snowden et devient très active après la publication des articles sur `XKeyscore` et `Bullrun` qui décrivent des attaques spécifiques contre l'infrastructure d'Internet. Les sujets abordés dans la liste :

- *brainstorming*
- critique de proposition pour l'augmentation de l'utilisation du chiffrement
- possibilité de création d'un nouveau WG pour des standards sur la sécurité
- discussions sur le modèle de menace et les réponses (dont la RFC 7258)
- discussions sur les procédures pour évaluer *privacy*

La réponse la + concrète : retournement de position en faveur du chiffrement des communications en ligne. Comprend le fait que des entreprises se mettent à chiffrer leur communication interne : par exemple *Google* se met à chiffrer les liens entre ses datacenters en réponse au programme de la NSA `muscular`. Au niveau des protocols Internet et Web, faire en sorte que la navigation soit chiffrée. Pour le W3C, le TAG met en évidence les étapes pour passer le Web en https. L'IAB a publié une déclaration sur la confidentialité, en encourageant le chiffrement à tous les niveaux disponibles pour les nouveaux protocoles (*no new cleartext*). Les fournisseurs de navigateurs ont annoncés qu'ils ne supporteront le nouveaux protocole http/2 qu'avec des communications authentifiée et chiffrées.

Les révélations Snowden montre que les attaques se font au niveau réseau, concernent donc aussi les Web APIs. Tentative de défendre l'idée de réstreindre les fonctionnalités du navigateur sensibles du point de vue de la *privacy* (p. ex les APIs pour accéder aux infos de localisation, camera et autres sensors), seulement aux pages chargées avec une connection sécurisée. Ont été débatues dans plusieurs W3C groupes. Motivations : mieux sécuriser les connections pour ces actions sensibles, mais aussi motiver les développeurs à déployer des connections plus sûres.

7 Tendances

7.1 Systématisation

Domaine de la sécurité le montre : expérience, conseils et aide améliore la substance et le sens des évaluations. Devrait être similaire dans domaine *privacy*.

IAB a tenté d'introduire RFC 6973 (*Privacy Considerations for Internet Protocols*) au moyen de tutoriels pour les directeurs des IETF WG et pour d'autres participants intéressés. Pour W3C, des individus ont créé des *checklists* similaires pour détecter les problèmes liés à la *privacy* dans les nouvelles Web APIs, ainsi que des conseils pour réduire les problèmes habituels de *privacy*, comme le *fingerprinting*

> Exposure of settings and characteristics of browsers can impact user privacy by allowing for browser fingerprinting. This document defines different types of fingerprinting, considers distinct levels of mitigation for the related privacy risks and provides guidance for Web specification authors on how to balance these concerns when designing new Web features.

Mais aussi d'autres propositions plus orientées "procédures", comme *Privacy Specification Assessment* : comme dans contexte gov. ou corp., décrit rôles, *workflows* et les délais pour identiifer et résoudre problèmes *privacy* au travers d'un cycle plus long pour concevoir et implémenter une spécification.

Systématisation continue, pas sans difficultés. Les standards du W3C sont développés de manière tellement décentralisées, il y a eu plus de succès pour obtenir des conseils et des évaluations informelles du *Privacy Interest Group* que de développer des procédures formelles pour la *privacy* tout au long de la conception de spécifications. IEFT a un *Security Directorate*, mais le *Privacy Directorate* a été abandonné pour manque d'activité.

7.2 Intégrer *privacy* et la sécurité

L'évaluation de *privacy* et sécurité se fait de plus en plus en même temps que le reste, plutôt que par des groupes ou des individus particuliers. Moins compliqué logistiquement. De plus, *privacy* et sécurité dépendent souvent des couches

inférieures, le travail d'identifier les problèmes est souvent similiares d'une spec à l'autre : rationalisation du travail.

Au W3C, deux groupes spécifiques : *Privacy Interest Group* et *Web Security Interest Group*. IAB remplace le *Secrutiy Program* et le *Privacy Program* avec un seul group discutant de :

- confidentialité
- résilience
- confiance

Contre la menace de surveillance massive, IETF a proposé de conduire des évaluation retro- et prospectivement.

7.3 Leadership

Établissement des standards au sein IETF et W3C : non hiérarchique. David Clark, in Tao of IETF :

> We reject kings, presidents and voting. We believe in rough consensus and running code.

Mais, *leadership* joue un rôle important dans dev. standards techniques, et significatif pour dev de considération *privacy* et sécurité. Les interviews avec des participants IETF identifie l'autorité des *Area Directors* et du IESG (Internet Engineering Steering Group), important pour *privacy* et sécurité dans les standards Internet. Les déclarations des orga *quasi-leadership* ont été importantes dans la réaction aux révélations Snowden.

Pourtant, des individus et des orgas se sont opposés à ce *leadership*, que ce soit pour la constitution des standards ou pour le rôle de gouvernance des standards. Des intérêts multiples et contradictoires. Ex: les vendeurs de "*middleboxes*" (proxies, cache, FAI : qui sont au milieux des connexions en réseaux) [voir : RFC 7663: IAB Workshop on Stack Evolution in a Middlebox Internet (SEMI) Report] se sont opposés aux projets de chiffrement de bout en bout, malgré le consensus apparent pour augmenter la confidentialité et le chiffrement. Des oppositions individuelles ont également été faites contre des choix par les groupes experts. Rappel la difficulté d'obtenir un consensus : il faut qu'il soit implémenté très largement pour avoir un impact.

8 Travail futur

Améliorer l'évalutation *privacy* pour Internet et Web standards dépend de bien des personnes et orgas. Mais comprendre comment ces thèmes sont mis en applications dans les standards fondammentaux d'Internet est crucial, alors que les orgas. tentent de réagir à l'attaque par les agences de renseignement contre les standards de sécurité.

Ce papier présente premiers résultats d'un travail en cours. But plus large : étudier, avec approche multi-modale et ethnographique, les pratiques qui affectent la *privacy* dans cadre d'orgas. constituant des standards pour Internet et Web : méthodes qualitatives dont entretiens semi-structurés pour 'attrapper' la diversité des participants et des parties prenantes; analyses textuelles automatiques et manuelles; et notes d'observation sur le terrain de la constitution des standards.

Le contexte de la fabrique des standards dans orgas. avec beaucoup de parties prenantes, diverses, et méthode pas trop formelles et hiérarchiques, défi pour l'implémentation de la *privacy* comme but de conception (*privacy-by-design*). Ce qui en ressort pourrait bien être utile dans des contextes similaires (orgas collaboratives sans structure hiérarchique), comme le logiciel libre (open source), les institutions de la gouvernance d'Internet (notez le pluriel), etc.

9 Remerciements

Un tas de beau monde, notamment pour "reviews", mais aussi le *U.S. Dept of Homeland Security...* :)