Ce que révèle le fichier des 500 000 patients qui a fuité

Que tout ça n'est pas très RGPD

48 • 38

INTERNET 9 MIN

Par Jean-Marc Manach
Le vendredi 26 février 2021 à 16:26

Signaler une erreur Offrir

Un fichier contenant les informations de 500 000 patients a fuité sur Internet et fait la une des médias ces derniers jours. S'il contient très peu de données médicales, il recèle des dizaines de milliers de mots de passe, numéros de téléphones, adresses postales et e-mails… « en clair ».

La fuite de données de santé concernant près de 500 000 patients, **révélée il y a quelques jours**, n'est pas si « *massive* » que cela. Après l'avoir analysée, nous relevons qu'elle touche 27 laboratoires d'analyses biologiques « *seulement* » (si l'on ose dire), situés dans le quart nord-ouest de la France (dont environ **350 000** Bretons).

Ce fichier recèle très peu de données médicales. La base de données contient un peu plus de 150 catégories, dont une cinquantaine portant sur des données des patients. Les plus sensibles sont les noms (y compris de jeune fille voire de conjoint au besoin), prénoms, adresses postales et électroniques, numéros de sécurité sociale, de téléphones fixes et de portables, ainsi que l'identifiant et le mot de passe utilisés pour accéder aux analyses biologiques.

payant. On note d'ailleurs qu'une bonne partie de ces catégories ne sont tout simplement pas renseignées dans le fichier.

Nous ainsi avons comptabilisé 489 838 numéros de sécurité sociale (ou **NIR**) et 478 882 personnes identifiées par leurs noms de famille, dont 268 983 femmes, 195 828 hommes, 13 478 qualifiées d'« *enfant* », 425 de « *bébé* » et 265 de « *sœur* ».

160 000 portables, 55 000 emails, 15 000 mots de passe

La base de données dénombre 270 569 numéros de téléphone fixe et 159 591 portables. À titre de comparaison, les 16 590 médecins fichés sont associés à 14 928 numéros de téléphones fixes et 1 971 portables. Le fichier répertorie également 55 738 adresses email uniques de patients et 337 de médecins.

Dans le « *top 10* » des domaines les plus présents que nous avons reconstitués, on trouve :

- 15 385 @gmail.com
- 14 418 @orange.fr
- 6 040 @hotmail.fr
- 5 579 @wanadoo.fr
- 3 530 @yahoo.fr
- 2 082 @sfr.fr
- 1 601 @hotmail.com
- 1 576 @free.fr
- 1 074 @live.fr
- 6 489 autres domaines

Mais 11,4 % seulement des patients dont des données ont fuité ont vu leur adresse email divulguée. Le fichier contient néanmoins des mots de passe, pour 14 997 d'entre eux (3 %), associés à leur « *identifiant SR* » (pour serveur de résultat, le site où ils étaient invités à récupérer les résultats de leurs analyses biologiques).

Comme on pouvait s'y attendre, nombreux sont les patients à utiliser des mots de passe similaires. Nous en avons décompté 11 443 uniques. À défaut de constituer un « *échantillon représentatif* » de la population, le fait qu'ils aient été utilisés par des personnes de tous âges et de toutes origines sociales dresse un aperçu instructif.

Sans surprise, on retrouve en effet le traditionnel et célèbre « *azertyuiop* », et 27 de ses déclinaisons (de type azerty1, AZERTY2, azerty22 ou encore azertAZERT). La base de données émanant pour la plupart de laboratoires bretons, y figurent également nombre de déclinaisons de breizh ou de Bretagne.

- **Choisir un bon mot de passe : les règles à connaître, les pièges à éviter**

Abr1bu$, M0u$71qu3, P4T0uch3

figurent dans le fichier) :

- Déclinaisons de dates : 01071981, 03sept69, 02.juju.19, 11septem, 11onzonz, 1948pierre ;
- Déclinaisons de prénoms, que l'on retrouve à foison : albert1, albert2, albert3, etc. ;
- Phrases ou expressions concaténées : jesuisgent, jesuismala, Je+suis+ne, jevaisbien, Jevisarenn, jevousaime, louisdefun, monmariage, Mot2passe, motdepasse, Pluscompli, prisedesan… et sans que l'on comprenne si la limitation à 10 caractères correspond à une « *fonctionnalité* » du logiciel, ou pas ;
- Combinaisons alphanumériques : 16abcdef, 13juju13, 17groseill, 1807mamie, 19gin19gin, 33COUcou, 44Loulou, 4896merde, belle2jour, bebe2019, faitchier9, mes3taupes, mes4loulou, mes5CHATS, mesamours3, mesange22, MesEnfant2, moncoeur21 ;
- Mots « *augmentés* » ou réécrits en ***leet speak*** : 3615Ulla, 37uD!4nt3$, Abr1bu$, bibi21CM, faitBeau?!., france1998, geishadu35, chiennedu22, idylle29, loverboy69, M0u$71qu3, P4T0uch3, rep0rt4g3, Rock1Rol.

On y trouve quelques récurrences étonnantes dans les mots les plus utilisés. Nous avons ainsi dénombré :

- 2 déclinaisons de bibiche, bichon, bidule, bidouille, cacahuete, cracotte, crapaud, framboise, frimousse, frisette, lapin, lapinou, mirador, nana, petitcoeur et toutouille,
- 3 de bichette, cactus, caline, calimero, cochon, diabolo, foufoune, grenouille, Looping, mama, maman, nounours et rocky,
- 4 de bonheur, cookie, espoir, ninouche, pompier, praline, reglisse, romeo, rose et tartine,
- 5 de bonjour, boubou, chaton, choupette, snoopy et tintin,
- 6 de cachou, chocolat, lamotte et pupuce,
- 7 de biscotte, chipie, chouchou, jetaime et minouche,
- 8 de canelle et mamour,
- 10 de noisette et princesse,
- 11 de louloute et soleil,
- 12 de doudou et gribouille,
- 16 de caramel,
- et 26 de loulou.

On trouve également plus de 4 300 mots de passe de type 124578AA, 124578AB, 124578AC, etc., itération laissant supposer qu'un ou plusieurs laboratoires auraient attribué un seul et même mot de passe à plusieurs patients (les identifiants, eux, étant différents). Un même mot de passe avait ainsi été attribué à plus de 800 patients différents.

L'analyse de ces mots de passe montre à tout le moins que les **recommandations**, en la matière, n'ont toujours pas franchi le plafond de verre et qu'il reste encore beaucoup à faire.

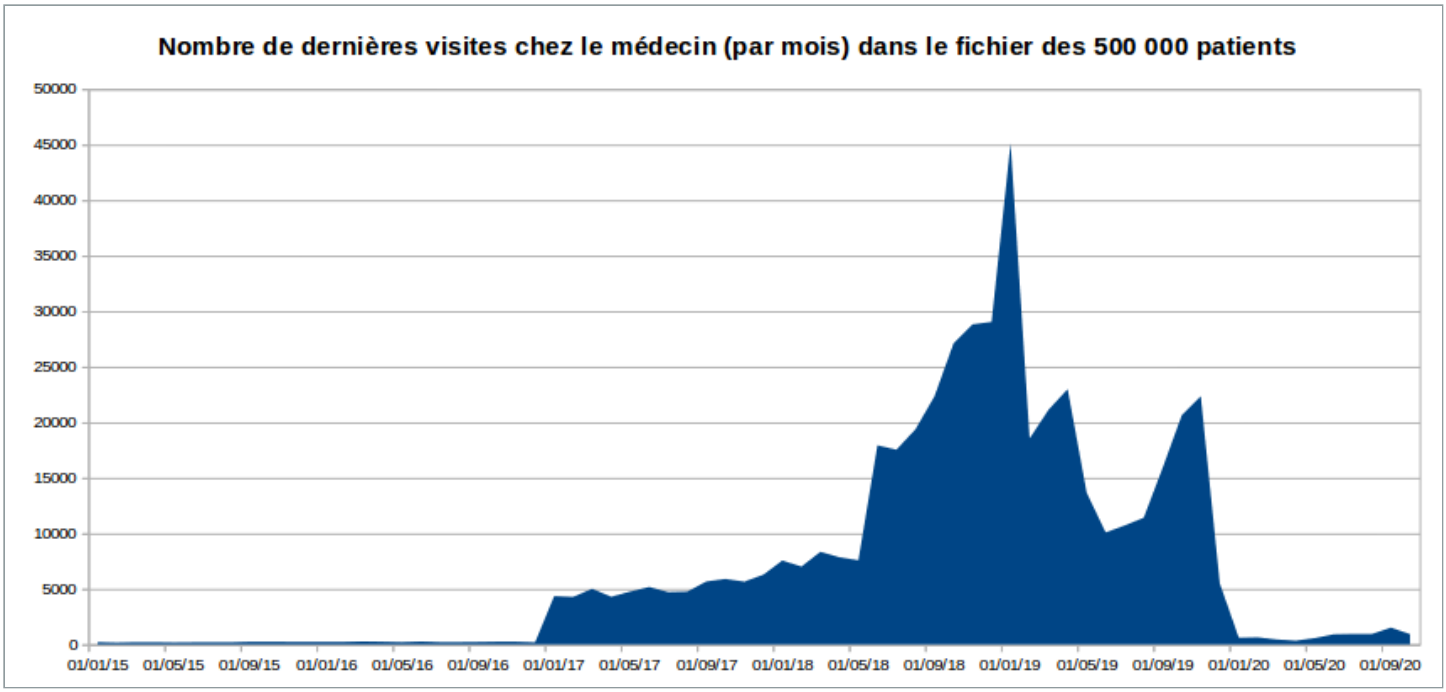
Un fichier « testé » en… mai 2018

Signe que la base de données repose sur un vieux logiciel, on y trouve une catégorie « *Consultation Minitel* ». Un indice révèle qu'il s'agirait vraisemblablement du logiciel Mega-Bus, comme l'avait reconnu son éditeur, Dedalus France. On y retrouve en effet l'identifiant d'un patient intitulé «

La rubrique « *dernière visite* » (chez le médecin) montre que le fichier a probablement lui aussi commencé à être renseigné à partir de 2018. Elle dénombre en effet 2 411 rendez-vous datant de 2015, 2 686 de 2016, 60 731 de 2017, mais 200 362 pour l'année 2018, 217 896 en 2019.

Seuls 7 780 rendez-vous datent de 2020, le dernier du 10 octobre, sans que l'on puisse vérifier si cette volumétrie en recul s'expliquerait par les effets du confinement, et/ou par un abandon croissant de Mega-Bus par les laboratoires.

Notre analyse montre une moyenne de 200 occurrences mensuelles pour les années 2015 et 2016, de 4 000 à 7 000 de janvier 2017 à mai 2018, puis de 17 000 à 28 000 de juin 2018 à fin 2019, laissant supposer que le déploiement du logiciel aurait bel et bien en lieu en mai 2018 :



Cette date est importante, d'une part parce que l'on sait **depuis 1999** que les mots de passe doivent non seulement être **hashés** et **salés**, mais également sécurisés au moyen d'une **fonction cryptographique**, avec l'objectif d'enrayer toute compromission, même si la base de données a fuité.

La CNIL a, à ce titre, publié **près d'une cinquantaine** de délibérations rappelant, depuis 2013, que « *les mots de passe ne doivent pas être stockés en clair en base de données* » et qu'elle « *recommande ainsi d'appliquer la fonction de hachage HMAC à clé secrète* ».

Le RGPD, adopté en avril 2016, est précisément entré en application en ce mois de mai 2018. En faisant le choix de recourir, à ce moment-là, à un logiciel ne sécurisant pas l'accès aux mots de passe, les responsables des traitements de données, tout comme le fournisseur du logiciel, ne respectaient pas le RGPD.

Ce dernier **dispose** en effet que les données à caractère personnel doivent être « *traitées de façon à garantir une sécurité appropriée* ». Il **précise** également que le montant des amendes administratives

Un des principaux logiciels de gestion des labos privés

L'Agence nationale de la sécurité des systèmes d'informations (ANSSI) a **indiqué** à l'AFP avoir identifié l'« *origine* » de la fuite des données de santé, et l'avoir signalée au Ministère des Solidarités et de la Santé en novembre 2020.

« *Les recommandations nécessaires ont été données par l'ANSSI pour traiter l'incident* », a-t-elle ajouté sans donner aucun détail supplémentaire. « *Présent au travers de son logiciel StarLab dans 400 laboratoires, Mega-Bus compte parmi les trois principaux éditeurs de logiciels de gestion pour les laboratoires privés d'analyse médicale en France* », précisait un **communiqué** publié à l'occasion de son rachat en 2009. « *A ce titre, cette société dispose d'une connaissance approfondie des spécificités liées aux besoins du secteur privé* ».

Son acquéreur, Medasys (depuis été rachetée par Dedalus France) était pour sa part présenté comme jouissant « *d'une position de leader national dans les domaines du dossier médical du patient et de la production de soins* ». Dans un **communiqué** laconique, l'entreprise vient d'expliquer que « *face à la gravité des sujets évoqués, Dedalus France est pleinement mobilisé et une enquête approfondie est en cours avec le support d'une équipe d'experts indépendants* ».

À Libération, l'entreprise avait émis l'hypothèse d'une fuite au moment des transferts vers les nouveaux logiciels, ou bien des problèmes de sécurisation des réseaux des laboratoires, se dédouanant au passage, sans mentionner le fait que les données n'auraient donc pas été chiffrées par son logiciel.

L'an passé, nous avions **révélé** que ce « *leader européen en matière de solutions logicielles de Santé* » avait licencié un lanceur d'alerte pour « *fautes graves* ». Il avait prévenu les autorités de ces problèmes de sécurité et découvert que « *n'importe qui pouvait accéder à l'extranet, depuis le web. Ce qui permettait notamment d'accéder aux tickets ouverts par les hôpitaux et laboratoires clients* ».

- **Un « leader européen » des données de santé licencie un lanceur d'alerte pour « faute grave »**

On devrait en apprendre plus d'ici peu. La Commission nationale Informatique et Libertés a en effet **lancé** mercredi des contrôles pour établir les manquements responsables de la fuite. Si l'ampleur de la fuite était vérifiée, l'affaire présenterait « *une gravité particulière* » au regard du nombre de victimes et de la sensibilité des informations médicales diffusées, a estimé Louis Dutheillet de Lamothe, secrétaire général de la CNIL.

Évoquant « *une violation de données d'une ampleur et d'une gravité particulièrement importante* », le gardien des données personnelles rappelle qu'il incombe aux organismes concernés de procéder à une notification auprès de la CNIL, dans les 72 heures. Mais également qu'ils ont en outre l'obligation d'informer individuellement les personnes concernées du fait que leurs données ont été compromises et publiées en ligne.

communication (OCLCTIC). « *On peut retrouver ce fichier à 7 endroits différents sur internet* », explique Damien Bancal, journaliste spécialiste de la cybersécurité, qui avait le premier **identifié** la fuite le 14 février sur son site Zataz.

Si vous voulez témoigner ou me contacter de façon sécurisée (voire anonyme), le mode d'emploi **se trouve par là**.

48 commentaires

fat - 26/02/21 à 16:30:15 **#1**

bibi21cm... Juste 😂

boogieplayer - 26/02/21 à 16:35:48 **#2**

Les mots de passe n'étaient pas chiffrés !? Ils étaient en clair ??!! O__O

Édité par boogieplayer le 26/02/2021 à 16:36

Arona - 26/02/21 à 16:39:36 **#3**

fat Je suis pas le seul à l'avoir vu celui-là!

xillibit - 26/02/21 à 16:44:41 **#4**

jesuisgent, jesuismala, Je+suis+ne, jevaisbien, Jevisarenn, jevousaime, louisdefun, monmariage, Mot2passe, motdepasse, Pluscompli, prisedesan

J'ai cru que c'était des fautes à la fin de certains mais on dirait que tous les mots de passe font 10 caractères maximum

Édité par xillibit le 26/02/2021 à 16:45

Sans intérêt - 26/02/21 à 17:05:18 **#5**

> cinquantaine de délibérations rappelant, depuis 2013, que « les mots de passe ne doivent pas être stockés en clair en base de données » et qu'elle « recommande ainsi d'appliquer la fonction de hachage HMAC à clé secrète ».

Je suis surpris que les mots de passes se doivent d'être *chiffrés*. À ma connaissance, aucune recommandation ne va dans ce sens. En revanche, le hashage de mots de passes, idéalement avec un sel spécifique à chaque utilisateur, afin d'éviter la reconnaissance de mots de passes identiques au sein même de la base de données, est en effet une recommandation systématique.

Z-os - 26/02/21 à 17:06:09 **#6**

xillibit Ou alors seuls les dix premiers caractères sont stockés et comparés. J'ai vu un site qui fonctionnait comme ça il y a longtemps.

manhack - 26/02/21 à 17:10:42 **#7**

Sans intérêt Comme indiqué, je faisais référence à bcrypt, une fonction de hachage basée sur l'algorithme de chiffrement Blowfish : **https://fr.wikipedia.org/wiki/Bcrypt**

grsbdl - 26/02/21 à 17:38:26 **#8**

> **boogieplayer a écrit :**
>
> Les mots de passe n'étaient pas chiffrés !? Ils étaient en clair ??!! O__O

Si tu savais le nombre de (souvent vieilles et dépassées) boites qui considèrent leur plateforme "sécurisée" par ce que personne d'extérieur n'est sensé connaître les urls des API critiques... Oui, ça fait bizarre quand le mec de la sécu te dit ça ^^. Pas de mot de passe, c'est open-bar, mais bon, il faudrait connaître l'IP et toute l'url, rholala ça n'arrivera jamais.
Le logiciel équipant une grande partie des offices hlm (OPAC, OPH) en faisait parti, par ex. La boite n'existe plus, mais pendant des années il y avait donc une énoooorme voie d'accès aux données, et pas de moyen de savoir si qqun s'était introduit dans le système.

Édité par grsbdl le 26/02/2021 à 17:40

alf.red - 26/02/21 à 17:46:42 **#9**

3615Ulla ^^
Mais bon, ça fait peine à voir toutes ces données en clair.

Merci pour l'analyse, vachement intéressant à lire :)

#10

Quelle horreur cette non sécurité totale...
Quand on voit des sociétés comme ça vendre des logiciels à des non-professionnels de l'informatique et dans des domaines ultra critiques, puis essayer de se dédouaner de toute responsabilité, c'est affligeant, ils n'ont aucun scrupule !
Je leur souhaite de perdre un maximum de clients suite à cette affaire lamentable.

Édité par Cronycs le 26/02/2021 à 17:53

Page 1 / 5

Votre commentaire

Connecté en tant que TheBigBug

Commentaire...

Envoyer