Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Département fédéral de la défense,
de la protection de la population et des sports DDPS

Office fédéral de la cybersécurité OFCS

7 mars 2024

Rapport concernant les analyses de données effectuées après la cyberattaque contre l'entreprise Xplain

Table des matières

1 Introduction 3
2 Contexte 4
3 Objectifs de l'analyse des données 5
4 Description des données publiées 5
5 Évaluation des données 6
6 Résultats de l'analyse des données 7
7 Analyse de l'écart entre le nombre de données publiées et le nombre de données à risque 9
7.1 Résultats des analyses 10
7.2 Évaluation du NCSC 11
8 Conclusion 11

1 Introduction

L’attaque par rançongiciel menée contre l’entreprise Xplain, un prestataire important de différents services fédéraux et de plusieurs cantons, a conduit à la fuite d’une grande quantité d’informations. Il s’agissait notamment de données importantes pour la sécurité et de données personnelles provenant notamment du domaine de la sécurité intérieure. Les cybercriminels ont publié les données volées sur le darknet, les rendant ainsi accessibles à des tiers. La divulgation de données confidentielles ou liées à la sécurité a de graves conséquences et génère une grande charge de travail. Des mesures d’urgence ont dû être définies et mises en œuvre afin de limiter les risques immédiats et d’informer les personnes concernées. Il a également fallu évaluer si des systèmes ou des banques de données de l’administration fédérale étaient compromis. De plus, des mesures ont dû être introduites pour que les systèmes puissent de nouveau être exploités. Une partie de ces travaux sont toujours en cours. L’incident a eu pour conséquence que des systèmes importants ne pouvaient plus être utilisés. De plus, des ressources humaines élevées ont été nécessaires pour le maîtriser. En sus de cet impact direct, la Confédération a subi un préjudice supplémentaire, car la confiance dans la sécurité des données qu’elle traite a été érodée. Il est difficile d’estimer quelles peuvent en être les répercussions. Cependant, il est certain que la Confédération doit tout entreprendre pour rétablir la confiance.

Dans ce but, l’une des premières étapes consiste à procéder à un examen complet de l’incident. Le Conseil fédéral a rapidement décidé de lancer cette analyse et a ordonné une enquête administrative le 23 août 2023. Cette dernière vise à déterminer si l’administration fédérale a rempli ses obligations de manière adéquate dans le cadre du choix de l’entreprise Xplain, des instructions qu’elle lui a transmises, de la surveillance qu’elle a exercée et de sa collaboration avec le prestataire. En outre, il s’agit d’identifier des mesures permettant d’empêcher un tel incident à l’avenir. L’enquête administrative doit s’achever d’ici la fin mars 2024. Le Conseil fédéral prendra ensuite connaissance de ses résultats et de ses recommandations et décidera de la suite à donner à ce dossier.

Le Centre national pour la cybersécurité (NCSC), aujourd'hui l'Office fédéral de la cybersécurité, a dirigé la gestion de l’incident, défini des mesures pour rétablir la sécurité des systèmes et effectué une analyse approfondie de l’ensemble des données publiées. Afin de contribuer au traitement de l’incident et d’assurer la plus grande transparence possible, il publie le présent rapport sur la procédure et les résultats de l’analyse des données. L’objectif est de donner un aperçu du type de données concernées et de présenter les défis liés à leur analyse.

La quantité de données publiées constitue également une question importante. Les cybercriminels ont d’abord annoncé avoir obtenu un total de 907 Go de données. Par la suite, ils n’en ont toutefois publié qu’environ 400 Go. Reste donc à savoir s’ils n’ont pas rendu certaines données publiques à dessein ou si leurs affirmations initiales étaient incorrectes. Il importe d’évaluer cette question afin de pouvoir estimer le risque que d’autres tentatives d’extorsion surviennent ou que les cybercriminels transmettent les données à des tiers. Le présent rapport montre comment cette question a été analysée et pourquoi le NCSC, en collaboration avec les services concernés et le Service de renseignement de la Confédération (SRC), est arrivé à la conclusion que le risque que les cybercriminels n’aient pas publié toutes les données volées à dessein est faible.

Il convient de souligner que le présent rapport ne procède pas à une évaluation du contenu des données et n’analyse pas non plus les raisons pour lesquelles certaines données ont pu faire l’objet d’une fuite. Ces questions seront clarifiées exclusivement au moyen des procédures en cours concernant l’incident, notamment dans le cadre de l’enquête administrative.

2 Contexte

Les rançongiciels à double extorsion sont une variante de l'attaque par rançongiciel lors de laquelle les cybercriminels, en plus de crypter les données, menacent également de publier les informations volées afin d'extorquer une rançon. Les victimes doivent donc payer non seulement pour décrypter leurs données, mais aussi pour éviter la divulgation d'informations sensibles. Le nombre de telles attaques a fortement augmenté, tant en Suisse qu'à l'étranger.
Fin mai 2023, un groupe de pirates informatiques se faisant appeler Play a mené une telle attaque contre la société Xplain, spécialisée dans les solutions informatiques liées à la sécurité intérieure. Il a volé de grandes quantités de données et a menacé de les rendre publiques. Comme l'entreprise, en accord avec les autorités de poursuite pénale et la Confédération, n'a pas cédé au chantage et n'a pas versé de rançon aux pirates, ces derniers ont publié les données dérobées sur le darknet le 14 juin 2023. Puisque l'administration fédérale fait partie des clients de Xplain, elle a aussi été touchée par ce cyberincident. Les services concernés ont mis en œuvre des mesures d'urgence afin de réduire les risques au minimum pour l'administration fédérale et les tiers lésés. Se fondant sur l'art. 12, al. 5, de l'ordonnance du 27 mai 2020 sur les cyberrisques (OPCy), le NCSC a géré les différents travaux opérationnels liés au traitement de l'incident et a effectué des analyses des données publiées sur le darknet.
Au début de l'incident, l'administration fédérale a fait face à plusieurs défis. Par exemple, on ne savait pas exactement quelles données avaient été dérobées ni quelle quantité cela représentait concrètement. De plus, on ignorait quelles unités étaient concernées et dans quelle mesure. En l'absence d'un tri et d'une analyse des données rendues publiques, il était impossible d'estimer les risques et l'ampleur potentielle des dommages pour la Confédération.
D'un point de vue pratique et opérationnel, l'accès aux données s'est déjà avéré très laborieux, car ces dernières ne pouvaient être téléchargées que très lentement à partir du darknet, et ce processus était régulièrement interrompu. Il a fallu plusieurs jours pour télécharger tout le stock de données et le préparer en vue de l'étape suivante. Le défi consistait maintenant à procéder à l'analyse forensique de ces données. Ce travail est particulièrement difficile pour les grandes quantités de données non structurées – comme dans le cas en question –, et ce, pour les raisons présentées ci-dessous.

- Diversité des formats de données : les données non structurées peuvent revêtir de nombreuses formes, comme des textes, des courriels, des images, des vidéos, des fichiers audio ou des formats techniques. Chacun de ces formats requiert des connaissances et des outils spéciaux pour être analysé. Le stock de données de Xplain contenait différents formats qui n'étaient pas directement lisibles : il a d'abord fallu les traiter et les préparer pour que leur contenu devienne compréhensible et interprétable.
- Pertinence des données : face à de grandes quantités de données, il est généralement difficile de distinguer les informations pertinentes de celles qui ne le sont pas. Dans le cas de Xplain, il a d'abord fallu déterminer quelles données étaient importantes pour l'administration fédérale et lesquelles ne l'étaient pas.
- Outils, ressources et savoir-faire : il est impossible d'analyser manuellement ou sans outils adéquats une grande quantité de données. Or les outils et les ressources nécessaires à l'analyse de données non structurées peuvent être coûteux, et leur utilisation requiert des connaissances techniques.

3 Objectifs de l'analyse des données

L'état-major de crise politico-stratégique « Fuite de données » (EMPS□F) [1] a chargé le NCSC de procéder à une évaluation des risques sur la base de l'analyse du stock de données.. Cette dernière a été réalisée en deux étapes : la première visait à trier et à catégoriser systématiquement tous les documents pertinents. La seconde a permis aux différents services de l'administration fédérale de rassembler les principaux résultats et de les comparer.

Le tri initial avait pour but de répondre aux questions ci-dessous.

- Quelles sont les unités administratives touchées et dans quelle ampleur ?
- À quels risques la Confédération est-elle exposée en raison de la publication ?
- Combien d'informations classifiées ont été rendues publiques ?
- Combien de documents contiennent des données personnelles ?
- Quels sont les autres enseignements tirés de l'analyse des données ?

Lors de la seconde étape, il s'agissait de déterminer les mesures requises, soit sur-le-champ, soit ultérieurement. Les données qui nécessitaient une action immédiate ont tout de suite été transmises aux services compétents. À la fin du tri, toutes les autres données pertinentes ont été remises pour examen aux unités administratives concernées. Seul le propriétaire des données ou l'unité administrative responsable peuvent procéder à une évaluation définitive du contenu et à une estimation des risques. Dans le cadre du traitement de l'incident, le NCSC a assuré le tri des informations en étroite collaboration avec les organisations concernées et d'autres autorités, notamment avec l'Office fédéral de la police (fedpol), les cantons, le réseau national de soutien aux enquêtes dans la lutte contre la criminalité informatique (NEDIK) et les exploitants d'infrastructures critiques.
En outre, le NCSC a coordonné les demandes émanant du public et, sur requête, effectué des recherches dans le stock de données afin que les services fédéraux compétents puissent indiquer si les personnes à l'origine des demandes étaient concernées ou non par la divulgation des données.

4 Description des données publiées

L'objet de l'analyse est le stock de données publié sur le darknet par Play, le groupe à l'origine de la cyberattaque. Dans le cadre de la gestion de l'incident de sécurité sous la direction du NCSC, les différents offices fédéraux et prestataires concernés ont collaboré étroitement, ce qui a permis de mettre à profit les synergies, d'utiliser les ressources judicieusement et de gagner un temps précieux.
Au total, 646 fichiers compressés au format RAR, d'environ 500 Mo chacun, et un fichier de 385 Mo ont été publiés sur le darknet. Le volume total des données compressées s'élevait à 339 Go. Les fichiers étaient chiffrés, le mot de passe permettant de les décrypter ayant également été publié par Play sur le site de la fuite. Après décompression, la quantité de données était de 431 Go, pour 146 623 fichiers et 19 863 répertoires.
Les indications sur les quantités servent à évaluer l'ampleur des données publiées. Il ne s'agit pas de chiffres précis ou définitifs, car les données quantitatives peuvent être interprétées de différentes manières et ne sont donc pas toujours comparables. En effet, la taille des données et les quantités indiquées varient selon la façon dont les fichiers ou les documents sont comptabilisés, par exemple sous forme compressée ou décompressée. Un fichier ZIP compromis peut contenir plusieurs autres fichiers, par exemple des courriels. Ces derniers peuvent à leur tour inclure des pièces jointes dans des formats comme Word ou PowerPoint. Ces documents Office peuvent également avoir d'autres objets tels que des images ou des

[1] https://www.ncsc.admin.ch/ncsc/fr/home/aktuell/im-fokus/2023/xplain_3.html

vidéos incrustés dans leur contenu. Dans le cadre de cette analyse, nous parlerons donc d'objets, lesquels pouvant se référer tant à des fichiers séparés qu'aux objets qu'ils contiennent.
La suite de la procédure avait pour but d'identifier la quantité totale de données pertinentes pour l'administration fédérale afin de rendre l'analyse et l'évaluation aussi efficientes que possible. Le stock de données contenait différentes copies de sauvegarde de serveurs de Xplain (deux fichiers, 260 Go) ainsi que d'autres données sans importance pour l'administration fédérale. Il s'agissait principalement de fichiers système, de programmes ou de composants standards utilisés par exemple pour le développement de logiciels. Ces éléments ont pu être exclus grâce à la comparaison automatisée des sommes de contrôle des stocks de données de référence (listes de hachage).
Les fichiers système et les fichiers de sauvegarde ont été vérifiés de manière automatisée et au moyen de recherches par mots-clés, puis exclus de la suite de l'analyse. Les données restantes, considérées comme importantes pour l'administration fédérale, comprenaient environ 65 000 documents et ont été soumises par la suite à un contrôle manuel et à un tri plus poussés. L'évaluation des données décrite dans le chap. 6 se réfère donc à la quantité de données pertinentes, quelque 65 000 documents, soit environ 5 % de l'ensemble du stock de données publié.

	Nombre d'objets	**Pourcentage**
Stock de données complet	1 295 862	100
Doublons	830 894	64
Données non pertinentes (sauvegardes, fichiers système, composants standards)	401 717	31
Données pertinentes : contrôle manuel et visionnement	**64 793**	**5**

Tableau 1 Vue d'ensemble du stock de données

5 Évaluation des données

Pour analyser de grandes quantités de données, on recourt à ce que l'on appelle des systèmes eDiscovery. Ces derniers permettent de préparer et de trier les stocks de données de manière efficiente et structurée. De plus, ils disposent de nombreuses fonctions de recherche qui permettent d'extraire rapidement les informations pertinentes à partir d'une grande quantité de données électroniques. En l'absence de tels instruments, la vérification des données prendrait beaucoup de temps et serait source d'erreurs. Dans le cadre de l'enquête sur Xplain, le NCSC a utilisé une solution eDiscovery interne à la Confédération, à savoir celle de l'Administration fédérale des contributions. Cette infrastructure professionnelle a permis de commencer les activités immédiatement et d'analyser et de visualiser les données de manière décentralisée et virtuelle au sein de l'administration fédérale.
Grâce à cette application eDiscovery, les quelque 65 000 documents ont été triés manuellement et classés selon une liste prédéfinie. En cas de doute concernant la classification ou l'évaluation, les documents étaient ensuite examinés par une deuxième personne (principe du double contrôle). Cela s'est produit pour environ 10 % des documents.
Le stock de données contenait également différents formats et artefacts numériques qui ne pouvaient pas être lus ni interprétés directement. Il a d'abord fallu les rendre compréhensibles, soit à l'aide d'un logiciel approprié, soit en utilisant des outils techniques développés à l'interne. De tels artefacts incluaient par exemple des fichiers journaux, des messages et des enregistrements d'erreurs, des sauvegardes de bases de données, du code logiciel, des configurations système, des formats codés d'images ou de documents, des images disques, des données chiffrées ou protégées par mot de passe, ainsi que des sauvegardes de différents formats ou produits.

Au début, 9 collaborateurs du NCSC ont participé à l'analyse des données. Ils ont en outre reçu le soutien de 27 volontaires provenant de dix offices fédéraux. Engagés pour la plupart à temps partiel, les 36 réviseurs ont consacré environ 1000 heures à ce travail. Les employés de la Confédération qui ont assisté temporairement le NCSC dans l'analyse des données dans le cadre de l'entraide administrative étaient considérés comme des auxiliaires du NCSC. Ils ont signé des accords de confidentialité concernant l'utilisation des informations issues de l'analyse des données.

6 Résultats de l'analyse des données

Le téléchargement et l'analyse de l'ensemble des données publiées nécessitant une charge de travail élevée, il manquait dans un premier temps une vue d'ensemble de qui était concerné et dans quelle mesure. L'objectif de l'analyse des données était d'y remédier.
Afin d'évaluer le degré d'impact et la responsabilité concernant les données publiées, une distinction a été opérée entre les *propriétaires des données* et les *unités administratives touchées*. Le propriétaire des données est celui qui a produit le document ou le stock de données, ou celui qui en porte la responsabilité. Il s'agit toujours d'un seul organe ou organisation. Cependant, des organisations peuvent être touchées par la publication de documents même si elles n'en sont pas les propriétaires. Cela peut notamment se produire si plusieurs organisations sont mentionnées dans un document. Un autre exemple concerne la réalisation d'un projet commun avec, dans le cas de Xplain, des participants issus de l'administration fédérale et des cantons. Les organes concernés ont été informés par le NCSC, soit directement, soit par l'entremise d'une organisation partenaire. Aucun autre détail à leur égard ne figure toutefois dans le présent rapport.
Le tableau 2 indique le nombre d'objets pertinents et leur pourcentage par rapport au total : la plupart des objets (plus de 70 %) appartiennent à l'entreprise Xplain. L'administration fédérale et les cantons possèdent respectivement environ 14 % et 10 % des données examinées. Les objets des organismes privés et des corps de police représentent chacun moins de 2 % de l'ensemble des données. Seuls quelques objets isolés ont pu être attribués aux entreprises liées à la Confédération ou au Ministère public de la Confédération (à chaque fois moins de 1 % du total).

Propriétaire des données	**Nombre d'objets**	**Pourcentage**
Xplain	47 413	73,03
Administration fédérale	9 040	13,92
Cantons	6 200	9,55
Organismes privés	955	1,47
Corps de police	944	1,45
Entreprises liées à la Confédération	355	0,55
Ministère public de la Confédération	16	0,02
Total	**64 923**	**100,00**

Tableau 2 Vue d'ensemble des propriétaires de données touchés

Le tableau 3 indique à quels départements appartiennent les 9040 objets pertinents de l'administration fédérale. Il montre ainsi que 95 % d'entre eux, soit l'immense majorité, proviennent des unités administratives du Département fédéral de justice et police (DFJP), à savoir de l'Office fédéral de la justice (OFJ), de fedpol, du Secrétariat d'État aux migrations (SEM) et du Centre de services informatiques (CSI-DFJP), le prestataire informatique interne au département. Avec un peu plus de 3 % des données, le Département fédéral de la défense, de la protection de la population et des sports (DDPS) est légèrement touché. L'impact sur les autres départements est marginal.
Les unités administratives touchées qui sont mentionnées ici en tant que propriétaires des données ne sont pas nécessairement clientes de Xplain et n'ont pas forcément de relation d'affaires avec cette entreprise. En effet, il se peut que leurs données aient été transmises à Xplain par l'intermédiaire de tiers, raison pour laquelle elles apparaissent dans le stock de données publié. Les tiers en question peuvent par exemple être d'autres unités administratives de l'administration fédérale ou des autorités cantonales avec lesquelles des données ont été échangées ou qui ont participé à des projets communs.

Propriétaire des données	**Nombre d'objets**	**Pourcentage**
DFJP (OFJ, fedpol, CSI-DFJP, SEM)	8 603	95,17
DDPS (PM, SRC)	306	3,38
DEFR (SECO)	69	0,76
DFF (OFDF, UPIC[2], OFCL, OFIT)	55	0,61
DFI (OFS)	6	0,07
DFAE	1	0,01
Total	**9 040**	**100,00**

Tableau 3 Propriétaires des données touchés au sein de l'administration fédérale

Dans le tableau 4, les 9040 objets de l'administration fédérale sont classés selon des critères de confidentialité (p. ex. données personnelles, informations techniques, informations classifiées, mots de passe). Au total, 5182 objets affichant un contenu sensible ont été recensés. Les données personnelles sont des objets qui contiennent des indications permettant d'identifier des personnes physiques (p. ex. nom, adresse électronique, numéro de téléphone, adresse postale). Un peu plus de la moitié des 9040 objets pertinents de l'administration fédérale contenaient des données personnelles, cette catégorie représente la plus grande part (plus de 90%) des objets sensibles.. Les informations techniques incluent des descriptions de systèmes informatiques ou leur documentation, des documents indiquant les exigences posées à des applications et des descriptions d'architectures. La catégorie des mots de passe comprend par exemple des données de connexion, des clés API ou des certificats cryptographiques. Les informations classifiées sont quant à elles des informations classées INTERNE, CONFIDENTIEL ou SECRET selon l'ordonnance du 4 juillet 2007 concernant la protection des informations (OPrI[3]).

Catégorie de données	**Nombre d'objets**	**Pourcentage**
Données personnelles	4 779	92,22
Informations techniques	278	5,36
Informations classifiées	121	2,34
Mots de passe	4	0,08
Total	**5 182**	**100,00**

Tableau 4 Données sensibles de l'administration fédérale

[2] L'Unité de pilotage informatique de la Confédération (UPIC) n'existe plus.

[3] RS **510.411**, https://www.fedlex.admin.ch/eli/cc/2007/414/fr

La répartition des 121 objets classifiés selon leur niveau de classification est la suivante : la plupart d'entre eux, soit 84 objets, sont classés INTERNE. Un tiers d'entre eux est classé CONFIDENTIEL. Aucun objet classé SECRET n'a été découvert.

Classification	**Nombre d'objets**	**Pourcentage**
INTERNE	84	69,42
CONFIDENTIEL	37	30,57
SECRET	0	0,00
Total	**121**	**100,00**

Tableau 5 Nombre de données classifiées selon leur niveau de classification

La répartition des objets sensibles par unité administrative montre que la plupart d'entre eux concernent des unités administratives du DFJP.

Propriétaire des données	**Données personnelles**		**Informations techniques**		**Informations classifiées**		**Mots de passe**	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**
DFJP (OFJ, fedpol, CSI-DFJP, SEM)	4 644	97,2	218	78,4	87	71,9	2	50
DDPS (PM, SRC)	76	1,5	22	7,9	24	19,8	0	0
DEFR (SECO)	31	0,7	26	9,4	9	7,5	2	50
DFF (OFDF, UPIC, OFCL, OFIT)	27	0,6	10	3,6	1	0,8	0	0
DFI (OFS)	0	0	2	0,7	0	0	0	0
DFAE	1	0	0	0	0	0	0	0
Total	**4 779**	**100,0**	**278**	**100,0**	**121**	**100,0**	**4**	**100,0**

Tableau 6 Données sensibles par département de l'administration fédérale

Les 3858 objets de l'administration fédérale qui ne sont pas considérés comme sensibles comprennent différents types de données comme des communications par courriel, la documentation de systèmes informatiques, du code source de logiciels, des documents liés à des projets, des descriptions d'erreurs, des demandes de support ou des captures d'écran d'applications.

7 Analyse de l'écart entre le nombre de données publiées et le nombre de données à risque

Sur son site web, le groupe Play a affirmé qu'il avait volé 907 Go [4] de données de Xplain et les avait ensuite publiés. Cependant, seuls quelque 400 Go de données ont été rendus publics sur le darknet. Cet écart n'a pas pu être expliqué et a fait craindre que d'autres données puissent être publiées ou que les criminels aient transmis une partie des données à des tiers.
En juillet 2023, l'EMPS□F a donc chargé le NCSC de procéder à une analyse de l'écart entre le nombre de données publiées et le nombre de données à risque de Xplain. À cette fin, il a fallu comparer de façon systématique les données stockées par Xplain au moment de l'attaque avec les données publiées. La société Xplain a transmis aux clients concernés de

[4] En février 2024, les auteurs ont corrigé la quantité de données à 600 Go.

l'administration fédérale les données structurées issues de la sauvegarde.
Les unités administratives ont soit mis ces données de sauvegarde à la disposition du NCSC pour que celui-ci puisse effectuer l'analyse, soit procédé elles-mêmes à l'analyse et informé le NCSC du résultat. Le NCSC a reçu des stocks de données de l'OFJ, du CSI-DFJP, du SEM et de l'Office fédéral de la douane et de la sécurité des frontières (OFDF) aux fins d'analyse. De leur côté, fedpol et armasuisse/Police militaire (PM) ont réalisé l'analyse par eux-mêmes.

7.1 Résultats des analyses

Les analyses ont donné les résultats suivants.

- fedpol : en comparant la valeur de hachage unique de chaque fichier de ses 70 Go de données restaurées avec l'ensemble des données publiées, fedpol a pu constater que seuls 39 % de ses données (52 468 fichiers) ont été publiés. Il manque parfois des fichiers et des sous-répertoires dans les cinq projets de clients rendus publics. Dans six répertoires, un écart a été constaté, avec à chaque fois un fichier non publié. Il est intéressant de noter que tous ces fichiers contenaient le mot « russisch » (*russe* en allemand) dans le nom du document (« InfoblattRussisch.docx »). Il semble que ces documents aient été délibérément supprimés (éventuellement de manière automatisée) de la fuite de données. Dans l'un des cinq principaux projets de clients, de nombreux sous-répertoires et fichiers n'ont pas été divulgués. Les noms de ces sous-répertoires se réfèrent au processus de développement logiciel (p. ex. « Realisierung », « Test », « Installation », « Dokumentation », « Abnahme », « Images », « Adressbücher », « Architektur », « Berechtigungen », « Codegruppen », « Domains PKI Mailer »). Il est difficile de savoir si Play a délibérément choisi de ne pas les publier.
- armasuisse/PM : le stock de données de 2,3 Go comprenait 1903 fichiers et 586 répertoires. L'analyse a été réalisée par la Base d'aide au commandement/milCERT et a donné les résultats suivants : les éléments non publiés incluaient notamment des offres détaillées (y c. documents d'appel d'offres), des factures et des contrats. Des documents relatifs à des tests de déploiement n'ont pas non plus été divulgués. Aucune systématique ne peut être identifiée. Il est possible que cela soit simplement une question technique et que certains répertoires ou fichiers n'ont pas été copiés durant le processus.
- Le NCSC a évalué les stocks de données de l'OFJ, du CSI□DFJP, du SEM et de l'OFDF, avec les résultats suivants.

Fichiers	**OFJ**	**CSI-DFJP**	**SEM**	**OFDF**
Total	26 719 (100 %)	16 961 (100 %)	178 (100 %)	30 757 (100 %)
Publiés	6 325 (24 %)	4 768 (28 %)	3 (2 %)	7 067 (23 %)
Non publiés	20 394 (76 %)	12 193 (72 %)	175 (98 %)	23 690 (77 %)

La plupart des données des unités administratives stockées chez Xplain n'ont pas été publiées. Un contrôle par échantillonnage a été réalisé sur les données non publiées afin d'examiner le contenu de ces documents. Il n'est pas possible d'identifier une systématique permettant d'expliquer pourquoi des données subtilisées n'auraient pas été rendues publiques. Il est probable que les données publiées aient été sélectionnées selon des critères techniques ou en raison des ressources limitées des criminels (temps, bande passante du réseau).

7.2 Évaluation du NCSC

Le fait que les pirates aient supprimé les documents nommés « InfoblattRussisch.docx » du stock de données de fedpol semble obéir à une certaine systématique, mais il indique plutôt qu'ils souhaitaient se protéger eux-mêmes. En effet, il est connu que d'autres groupes cybercriminels russes évitent de faire référence à la Russie pour ne pas risquer d'attirer l'attention de leurs propres autorités. Le document en question est une fiche d'information en russe pour les personnes arrêtées (droits et obligations, procédure, etc.) sans contenu sensible. Il est présent dans différentes autres langues parmi les données publiées. Une analyse plus poussée a montré que le fichier « InfoblattRussisch.docx » n'a été retiré du stock de données que s'il n'était pas contenu dans une archive (p. ex. fichier .zip) et qu'il était donc facile à trouver. Cette manière de procéder indique que les criminels ont réalisé un examen superficiel du stock de données sans en parcourir le contenu.
Le NCSC ne dispose d'aucune information suggérant que Play aurait retenu des données spécifiques lors de la publication des données volées aux autres entreprises ou organisations. De même, un examen superficiel et par échantillonnage des différentes fuites publiées par Play a révélé que le groupe avait déjà fourni dans d'autres cas des indications erronées ou imprécises concernant le volume des données dérobées.
En d'autres termes, l'analyse des écarts n'a pas permis d'établir une intention délibérée de conserver des données. Le NCSC estime donc que le risque de nouvelles divulgations est faible. En outre, selon les résultats de l'analyse, rien n'indique que les criminels aient transmis une partie des données à des tiers.

8 Conclusion

Lorsqu'une cyberattaque conduit à une fuite de données, les victimes doivent se demander non seulement comment rétablir l'exploitation opérationnelle, mais aussi quelles sont les répercussions de la perte des données. Ces questions sont particulièrement importantes pour la Confédération, car cette dernière traite les données concernées sur mandat public. Leur pertinence était d'autant plus grande dans le cas de l'attaque contre Xplain, puisqu'il s'agissait d'un prestataire informatique actif dans le domaine de la sécurité intérieure et que des données très sensibles étaient donc concernées.
Le rapport a montré qu'un travail d'analyse considérable est déjà nécessaire pour déterminer quelles données ont été volées, avec quelle ampleur et à qui elles appartiennent. Les données non structurées doivent d'abord être préparées et rendues lisibles au moyen d'instruments appropriés. Ensuite, il n'y a d'autre choix que de trier et de catégoriser à la main les données jugées pertinentes. Face à de grandes quantités de données, il s'agit d'un travail considérable qui demande beaucoup de temps et de personnel.
Cependant, il est impératif de procéder à une analyse complète des données afin que l'incident puisse réellement être évalué. Si ce processus n'est pas mené à bien, il est impossible de savoir qui est réellement concerné et dans quelle mesure. De plus, le risque que des données importantes soient publiées sans que les personnes concernées en aient conscience serait trop élevé. En outre, une analyse complète des données constitue une condition importante pour que l'incident puisse être traité. Les conclusions pertinentes à cet égard seront reprises dans l'enquête administrative ordonnée.
Comme souvent en situation de crise, l'analyse des données a requis beaucoup d'improvisation. Le NCSC a toutefois pu constater qu'il a été possible de mobiliser assez rapidement les infrastructures et le personnel nécessaires au sein de la Confédération. Grâce à l'analyse des données et à une collaboration étroite avec toutes les parties concernées, il a pu fournir aux décideurs politiques et au public des informations aussi transparentes que possible.

Malgré toutes les mesures de précaution, il faut malheureusement s’attendre à ce que des incidents impliquant des données de la Confédération se produisent encore à l’avenir. Il convient de clarifier les compétences au sein de l’administration fédérale et de développer les capacités opérationnelles afin d’analyser les données rapidement et de manière exhaustive.