Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
CONSEIL DE L'INNOVATION

Grand Défi cyber-sécurité

« Automatisation de la cyber-sécurité »

Feuille de route – William Lecat

I. Rappel général sur l'initiative des Grands Défis

L'initiative des Grands Défis est issue, à l'origine, d'une recommandation du rapport Villani sur l'intelligence artificielle. Elle a été reprise et élargie au-delà de l'IA par le Conseil de l'Innovation (constitué de 7 ministres et 7 personnalités qualifiées, co-présidé par Bruno Lemaire et Frédérique Vidal). Le principe reste le même et s'inspire de ce que peut faire la DARPA (l'agence d'innovation de défense américaine). Il s'agit de sélectionner une thématique avec un fort enjeu économique et sociétal, de nommer un directeur de programme sur le sujet, de le doter d'un budget de 30M€, d'un mandat de principe de trois ans et d'une grande autonomie pour avoir un impact sur la thématique (en faisant émerger des avantages technologiques, économiques, voire stratégiques). Quatre Grands Défis ont déjà été lancés : IA de confiance, diagnostic médical et IA, automatisation de la cyber-sécurité et bioproduction médicale.

II. Notions de cyber-sécurité

La cyber-sécurité est un domaine très vaste aux contours parfois un peu flous. La description qui suit n'a pas pour vocation d'être exacte ou particulièrement précise, mais plutôt de donner des notions de base pour fournir un canevas d'analyse aux lecteurs peu familiers avec le sujet.

Le modèle standard de la cyber-sécurité est centré sur les trois concepts suivants : confidentialité, intégrité et accessibilité. Pour expliciter un peu plus ces aspects, l'exemple de l'email est parlant et simple. On souhaite pouvoir échanger des emails sans que tout Internet puisse les lire (dans l'idéal juste l'expéditeur et les destinataires devraient avoir cette capacité). De même, on attend de ce service d'emails d'avoir une garantie sur l'authenticité du mail (i.e. qu'il n'ait pas été modifié en chemin) et de l'expéditeur (ce qui n'est souvent pas le cas en réalité). Enfin, on doit pouvoir avoir accès à sa messagerie pour pouvoir utiliser ce service.

Ce modèle peut être appliqué de manière immédiate aux données, mais a toute sa place au niveau des équipements et des réseaux. Les réseaux contiennent les équipements qui contiennent eux-mêmes les données. Il y a donc plusieurs niveaux d'observation où la sécurité peut s'appliquer de manière complémentaire. Cette approche technique doit être complétée par son pendant organisationnel au niveau de l'information, de l'utilisateur et de l'organisation.

Enfin, la mise en œuvre de cette cyber-sécurité peut être comprise en trois phases, qui se répètent éventuellement pour former un cycle : la phase amont de maitrise du risque (analyse de risque, conception, protection, parfois qualification, etc.), la phase d'opération (détection, supervision) et la phase de remédiation (maintien en condition de sécurité avec les mises à jour, restauration, réponse à incident, etc.).

III. Organisation, constats et grands principes

a. Démarche préalable

Pour permettre la rédaction d'une feuille de route à trois ans, basée sur des constats issus du terrain et en complément de la lettre de mission originale, une démarche de rencontre d'un maximum d'acteurs de l'écosystème a été opérée (et se poursuivra probablement au moins un temps), permettant de bien identifier le positionnement de chacun et les ambitions existantes. Cela a ainsi permis de rencontrer plus d'une centaine d'acteurs et de nourrir la feuille de route, avec une vision la plus globale possible. L'intérêt s'est porté sur tous les travaux en cyber-sécurité (ou pouvant s'y appliquer) pour recueillir un maximum d'idées. Les seules guidelines fixées au préalable ont été de se concentrer sur des projets de développement logiciel (par opposition au « hardware ») de produits (par opposition aux services, même si le développement de produits ayant vocation à être intégrés dans une offre de service reste dans le périmètre).

b. Constats

Cette démarche a permis d'établir plusieurs constats regroupés dans 5 axes. Les trois premiers axes, correspondant à des thématiques distinctes et qualifiés de « verticaux », sont : « le dynamisme des réseaux », « les objets connectés (IoT) » et « la protection des petites structures contre la cybercriminalité ». Les deux autres axes, qualifiés de transverses car ils touchent plusieurs axes verticaux et peuvent dépasser le cadre du Grand Défi, sont : « l'amorçage en cyber-sécurité » et « la problématique des données cyber ». L'approche globale est donc résumée de la manière suivante :

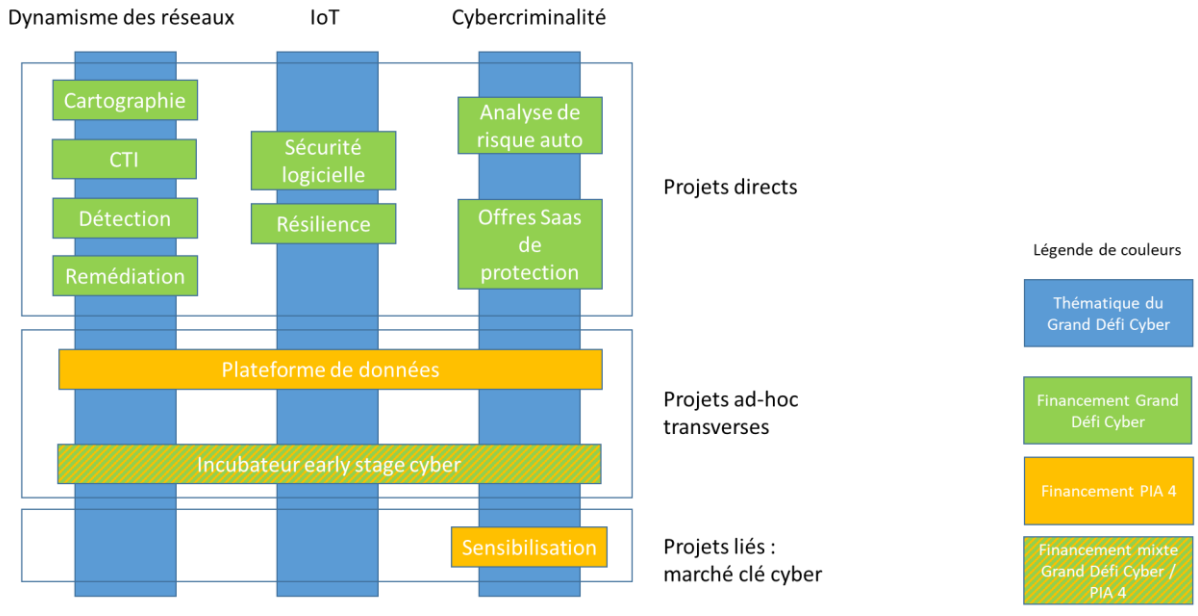

Figure 1 : approche globale du Grand Défi Cyber

c. Positionnement des financements

Dans le cadre du défi cyber, l'idée est d'investir sur des approches innovantes pouvant donner à la France et à l'Europe une avance technologique et un avantage compétitif. Il s'agit bien de se positionner sur des aspects de « rupture », faisant émerger de nouveaux types de technologies ou d'acteurs. On se focalise en particulier sur les aspects ambitieux trouvant difficilement des financements portés uniquement par des acteurs privés (car ils sont trop risqués techniquement, trop long terme, etc.) ou nécessitant une dynamique globale et une fédération de l'écosystème.

d. Sélection des projets

Le mode de sélection des entreprises ou des projets reste très ouvert. Néanmoins, la contrainte de temps (la durée du programme est de 3 ans) et la vitesse d'évolution du domaine sont très importantes. Dans ce contexte, et en cohérence avec la démarche actuelle, le mode de sélection le plus simple sera d'identifier directement les projets (appelé « projets directs » dans la suite), les acteurs pertinents et leurs besoins de financement au travers des discussions et des rencontres de l'écosystème. Cela n'exclut pas la possibilité de mise en place d'appels à projets thématiques pour faire émerger ou identifier de nouveaux acteurs. Dans tous les cas, les acteurs intéressés doivent contacter le directeur de programme pour présenter leur projet éventuel, garantissant ainsi à tous la possibilité d'entrer dans le processus de sélection du Grand Défi. Une grille d'éligibilité et de sélection sera établie pour permettre une sélection transparente et objective.

e. Types de financements

Là aussi, le sujet est très ouvert. Il reste quand même un certain nombre de contraintes dues à la simplicité (ou difficulté) de mise en œuvre administrative. Le plus simple et le plus efficace sera de soutenir les acteurs *via* des subventions R&D. Si cela se justifie par une valeur ajoutée importante, il sera aussi envisageable, au cas par cas, d'apporter des financements en fonds propres (ou quasi fonds propres) pour les acteurs ne pouvant bénéficier d'un soutien en subventions. Néanmoins, quand cela sera possible, la priorité sera mise, si le financement nécessaire est en fonds propres, sur l'orientation vers des outils existants utilisables conjointement au Grand Défi comme le fonds opéré par Bpifrance « Digital Venture », ou les fonds PIA « ambition amorçage angels » et « French Tech Seed ».

f. Constitution des projets

Les projets se constitueront le plus possible autour d'un modèle composé de deux types d'acteurs. Tout d'abord, un acteur (ou un groupe d'acteurs) réalisant le projet innovant. Ensuite, un deuxième acteur qui amènera un cas d'usage concret. Cela permettra de garantir que le projet répond réellement et directement à une problématique concrète, facilitant l'accès au marché en fin de développement. Cet acteur devra permettre des expérimentations en boucle courte, rendant ainsi possible un cycle court complet (développement -> déploiement -> exploitation -> retours). Enfin, les acteurs réalisant le projet devront pouvoir justifier des financements complémentaires pour mener à bien le projet. En effet, dans la majorité des cas, le Grand Défi apportera un financement sous forme de subvention R&D, ce qui a plusieurs implications. D'abord, cette subvention n'adressera que la partie R&D du projet, alors que la partie commerciale sera tout aussi importante pour le succès du projet. Ensuite, en conformité des règles fixées par la Commission européenne, sur un projet de développement logiciel

en cyber-sécurité, le montant de subventions ne dépassera pas, en général, 50 % de la charge R&D. Cela laisse donc une partie importante du projet à financer (en auto financement ou *via* un co-financeur, bancaire par exemple) garantissant le partage du risque.

g. Jalonnement des projets

Les projets s'organiseront dans une suite de jalons relativement courts (6 à 12 mois) : technologiques, de démonstration et éventuellement commerciaux. L'idée est de pouvoir, à chacune de ces étapes, réévaluer l'avancée des travaux, l'atteinte des objectifs et les perspectives de développement à venir. Cela permettra l'arrêt éventuel des projets qui sous-performeront, ainsi que la réorientation de certains projets si cela est nécessaire. De plus, chaque projet se déroulera en deux phases sur un modèle d'entonnoir : une première phase où le financement sera accordé à différents projets (appelée « tranche – financement spécifique ») et une seconde, optionnelle, où des financements complémentaires seront accordés à certains des projets déjà soutenus, dans une logique de mise en concurrence, pour approfondir les premiers développements. Cela permettra de pousser les projets les plus performants plus loin sans avoir à prévoir lesquels dès le lancement.

h. Recherche académique

Afin de favoriser des aspects plus amonts dans les projets, l'opportunité de thèses CIFRE sera systématiquement explorée dans la constitution des projets. Cette démarche sera complémentaire à la participation directe de certains laboratoires dans les projets et pourra concerner des périodes dépassant l'horizon du Grand défi (par exemple trois ans pour une thèse CIFRE).

i. Défragmentation du marché

Les projets auront vocation à être des « macro projets » regroupant d'autres sous-projets cohérents (mais restant indépendants). L'offre française en cyber-sécurité est très fragmentée (surtout vue depuis l'étranger). En effet, beaucoup d'acteurs offrent des solutions pertinentes mais sur un secteur restreint de la cyber-sécurité, rendant ainsi difficile leur acquisition seule. Chaque « macro projet » du Grand Défi aura, entre autres, pour but de regrouper des acteurs aux projets complémentaires. Ainsi, l'objectif de chaque acteur sera double. Tout d'abord, faire avancer son projet individuel de manière ambitieuse. Ensuite, rendre sa solution interopérable avec les autres solutions du « macro projet » pour permettre de faire émerger une solution globale cohérente, plus visible, facilement intégrable et complète (sur son périmètre).

j. Angles de l'automatisation

Les termes présents dans l'intitulé du Grand Défi ont volontairement été laissés sans définition plus précise pour éviter l'autocensure de la part d'acteurs potentiellement pertinents. L'objectif est d'entendre et de discuter un maximum d'idées dans le domaine de la cyber-sécurité ou pouvant s'y appliquer. L'automatisation est presque toujours un facteur des projets innovants dans ce domaine. Pour le moment trois aspects semblent se dégager du terme « automatisation » mais la liste peut encore s'allonger : l'automatisation pour traiter plus de données (passage à l'échelle des solutions), l'automatisation pour mieux traiter les données (amélioration de l'efficience des outils) et l'automatisation pour rendre la cyber-sécurité transparente à l'usage (démocratisation de la cyber-sécurité).

IV. Thématiques identifiées

a. *Axe vertical 1 :* le dynamisme des réseaux

Les réseaux informatiques et industriels subissent, depuis plusieurs années déjà, deux mutations fondamentales. Tout d'abord, une convergence des réseaux IT (informatique « standard ») et OT (systèmes industriels) s'opère rapidement. Cette mutation a été rendue possible, voire nécessaire, par la prédominance accrue du protocole IP pour les communications numériques, en conjonction avec la numérisation rapide des systèmes industriels. Cela a l'avantage d'uniformiser les réseaux, de rationaliser les équipements et les compétences et de faciliter la maintenance tout en apportant à l'OT une grande quantité de fonctionnalités développées, à l'origine, pour l'IT. L'inconvénient associé est, bien sûr, l'arrivée sur le « monde OT » de toutes les menaces existantes dans l'IT. La deuxième mutation concerne le dynamisme de ces réseaux. En effet, d'une part, la généralisation de la virtualisation permet de reconfigurer les réseaux « à chaud » et de manière logicielle (comme, par exemple, le prévoit la 5G ou le « Software Defined Network » (SDN)) et, d'autre part, l'arrivée massive des IoT, qui rend potentiellement mobiles beaucoup d'élément de ces réseaux. En complément, l'adoption généralisée du « cloud » (qui utilise largement la virtualisation et contribue à favoriser la mobilité) accentue largement ce phénomène. En ajoutant à cela les smartphones, tablettes et ordinateurs portables, on constate que ces réseaux sont extrêmement dynamiques et mouvants. Cela induit d'abord un problème évident de supervision, mais, au-delà de cet aspect, c'est tout le modèle traditionnel du cycle de la sécurité qui est remis en question. En effet, une démarche standard d'analyse de risque dont découlerait un plan d'action puis des audits annuels pour vérifier son application est rendue obsolète par ces changements. L'analyse de risque qu'on espérait pertinente pour plusieurs années l'est à présent pour quelques minutes. Le modèle doit donc être repensé pour gérer cette nouvelle situation qui représente une tendance de fond. L'analyse de risque générale reste très pertinente, mais doit être complétée par un modèle adapté à des évolutions de court terme et rapides.

Trois composantes essentielles doivent être rapprochées et mises en relation en temps réel : la « Cyber Threat Intelligence » (CTI), la modélisation du risque et la supervision. D'une part, la modélisation du risque doit bénéficier d'informations en temps réel concernant les évolutions du réseau (potentiellement *via* la supervision) pour mettre à jour les modèles de menace. Ces modèles doivent également évoluer en temps réel grâce à la contextualisation apportée par la CTI. Ainsi, l'analyse des menaces pourra rester pertinente sur la durée et permettra une véritable catégorisation automatique des alertes de supervision alors que c'est peu le cas aujourd'hui. Enfin, les centres de supervision de sécurité (SOC) ne passeront pas à l'échelle sans se doter de capacités (simples au départ) de réaction automatisées pour initier les levées de doute sur les alertes *via* des investigations complémentaires. Cette automatisation passe nécessairement par l'orchestration de tous les produits de sécurité. Malheureusement, nous avons aujourd'hui entre 3 et 5 ans de retard (sur les Etats-Unis) concernant ce sujet. En effet, de nombreux grands groupes internationaux (américains en particulier) ont massivement racheté des sociétés pour acquérir rapidement ce type de nouvelle technologie appelée SOAR (« Security Orchestration, Automation and Response »). Le premier a été FireEye, rachetant Invotas en 2016, suivi d'IBM avec le rachat de Resilient Systems en 2016 également. En 2017, Microsoft a payé 100 M$ pour Hexadite et Rapid7 la moitié pour Komand. En 2018, Splunk a fait l'acquisition de Phantom (une société au passage financée par In-Q-Tel, indication d'une certaine tendance). Enfin, Palo Alto a procédé au rachat de Demisto en 2019. Il n'est pas encore trop tard pour adresser cette

technologie de manière innovante en France, le domaine restant très jeune et les choix technologiques n'ayant pas encore été validés par l'expérience. La clé de cette orchestration passera par une exploitation des composants de sécurité mis à disposition nativement sur les systèmes. Le comble aujourd'hui est de constater que ces outils sont massivement détournés par les attaquants (cette stratégie a été nommée « Living off the land » ou LotL) mais pas totalement exploités par les défenseurs.

A l'autre bout de la chaine, la « (Cyber) Threat Intel(ligence) » nécessite également un effort important de consolidation. Aujourd'hui beaucoup d'acteurs français se positionnent sur ce créneau, mais l'émergence d'un acteur majeur (i.e. pouvant concurrencer les plus grands acteurs étrangers) n'est toujours pas garantie et se heurte également à la problématique des données de cyber-sécurité (cf. Axe transverse 2).

Objectifs :
Adresser les problèmes de sécurité dus au nouveau dynamisme des réseaux sur la base de l'émergence :
- *D'une offre de cartographie*
- *D'acteurs de cyber threat intelligence (CTI)*
- *D'un nouveau modèle complémentaire de modélisation du risque*
- *De capacités rapprochées et automatiques d'orchestration, de détection et de remédiation*

Typologie d'acteurs :
- *Start-up et PME innovantes*
- *Laboratoires de recherche*
- *Partenaires industriels, utilisateurs finaux*

Calendrier :
- *Lancement des projets de septembre 2020 à avril 2021*
- *Tranche de financement spécifique jusqu'en février 2022*
- *Tranche de financement concurrentiel en 2022*

b. *Axe vertical 2 :* les objets connectés (IoT)

L'IoT ou Internet des Objets est un véritable « buzzword » que l'on entend depuis longtemps déjà. L'idée de disposer de capteurs sans fils n'est pas nouvelle et n'a pas présenté jusqu'à maintenant une révolution. Le changement aujourd'hui est double. Tout d'abord, tous les « objets » qui étaient déjà connectés par le passé le sont de plus en plus en utilisant le protocole IP (Internet Protocol) et en intégrant des composants logiciels issus de l'IT, d'où la notion d' « Internet » des objets. Ensuite, nous connectons (à Internet) de nouveaux objets (télévision, réfrigérateur, ampoules, etc.) pour fournir de nouveaux services, d'où la notion d' « objet connecté » qui signifie véritablement « objet connecté à Internet ». Il y a donc un véritable phénomène de convergence (du précédemment connecté hors IP et du précédemment non-connecté) vers la technologie de l'Internet. Dans la suite, nous ne nous intéresserons donc qu'aux objets connectés par le protocole IP qui seront globalement appelés « objets connectés » ou « IoT ».

Du point de vue du réseau, cela signifie une triple nouveauté : de nouveaux types d'éléments, dans de nouveaux contextes et en nombre beaucoup plus importants. Premièrement, les nouveaux types d'éléments induisent en particulier une interaction avec le monde physique, ce qui était auparavant

réservé au monde industriel (Operational technology ou OT par opposition à l'Information technology ou IT). Cela peut concerner une serrure de porte, un système de freinage ou une ampoule... Gardons ce dernier exemple pour explorer les deux autres aspects de nouveauté. Deuxièmement, le contexte est très souvent aussi nouveau, puisque cette ampoule connectée peut se retrouver partout : dans une cuisine, dans une usine, dans un véhicule. De plus, ce contexte peut inclure un grand nombre d'autres objets eux aussi connectés et embarqués dans la même plateforme (potentiellement en mouvement dans le cas d'un véhicule par exemple). Enfin, troisièmement, pour avoir une notion du nombre « beaucoup plus important » d'objets inclus dans le réseau, il suffit d'imaginer que toutes les ampoules du monde deviennent progressivement connectées (ainsi que toutes les portes, les fenêtres, les panneaux de signalisation routière, etc.). Jusqu'à aujourd'hui, des réseaux comportant de l'ordre de 100 000 postes utilisateurs appartenaient à la classe des « gros » réseaux. On imagine bien que ces « gros » réseaux comporteront demain de l'ordre du million, de la dizaine de millions ou plus d'éléments connectés en interaction avec des utilisateurs (capteurs de tout type, actionneurs, affichages, etc.). La croissance déjà rapide des réseaux IP est en train de changer d'échelle. Rappelons-nous que le passage de la version 4 à la version 6 de l'IP sur Internet, principalement motivé par la pénurie d'adresses IP (chaque utilisateur sur le réseau doit avoir une adresse IP propre et la protocole IP v4 n'autorisait « que » 4 milliards d'adresses différentes), dure depuis plus de 10 ans et n'est toujours pas achevé. Cela donne une idée de la croissance passée, qui n'a pourtant rien à voir avec celle à venir. Du point de vue de la cyber-sécurité, ces trois aspects de nouveauté sont trois challenges que la technologie actuelle ne permet pas d'adresser.

Tout d'abord, l'objet lui-même amène un challenge de sécurité. Des contraintes de consommation, d'encombrement, de capacité de calcul et de coût (un objet prévu à plusieurs millions d'exemplaires cherche à intégrer des composants à bas coûts) ont favorisé l'absence totale de sécurité. Des démonstrations dramatiques du besoin de sécurité à ce niveau et les enseignements issus du monde de l'IT permettent une amélioration rapide dans ce domaine tout en se heurtant aux contraintes cités précédemment. Malgré des exemples d'attaques mondiales, comme ce fut le cas avec le botnet Mirai qui se servait des caméra IP peu sécurisées pour opérer des vagues massives de déni de service distribué, il n'en reste pas moins que la sécurité sur un objet connecté demande potentiellement des composants relativement plus chers. Aujourd'hui, les progrès des composants électroniques permettent de s'affranchir des limites de capacité de calcul et de consommation. Il demeure néanmoins une problématique de coût (relatif), qui freine l'adoption mais surtout le développement d'innovation de sécurité à ce niveau. La problématique du coût, en elle-même, sera résolue par une combinaison de trois facteurs : obligation d'intégrer de la sécurité, demande client pour la sécurité et baisse des coûts des technologies de sécurité. Le premier aspect peut se traiter par la réglementation, le deuxième par la sensibilisation des utilisateurs (éventuellement indirecte, *via* l'observation d'attaques ayant des conséquences importantes) et le troisième par la R&D de solution innovante pour élever le niveau de maturité du sujet.

Ensuite, les plateformes autonomes embarquant massivement des objets connectés représentent également un challenge de sécurité. Il existe souvent des contraintes d'encombrement (empêchant d'embarquer un SOC ou un opérateur de niveau 1) et de bande passante (empêchant d'opérer un SOC à distance sur le modèle actuel). C'est par exemple le cas des véhicules connectés, mais aussi de beaucoup de plateformes temps réel. Ce type de plateformes doit être capable de prétraiter certains événements de sécurité, voire d'y réagir de manière autonome (pour les situation temporellement critiques) pour ne remonter au centre de supervision que l'essentiel (et ainsi éviter de saturer les faibles bandes passantes). Il y a donc ici un modèle à définir, ce champ restant très largement inexploré aujourd'hui.

Enfin, le troisième challenge sécuritaire, au niveau du SOC lui-même, fait partie d'une problématique plus large qui fera partie d'un axe d'effort à part entière : « le dynamisme des réseaux ».

L'étendu du marché des IoT et les challenges qu'il représente en cyber-sécurité en font, à court terme, un domaine de compétitivité essentiel pour les acteurs du domaine.

Objectifs :
Faire émerger une offre de sécurité logicielle de bout en bout :
- De l'aide au développement à l'analyse de firmware en passant par l'analyse de code source et de binaires
- Permettre la validation automatisée de la supply chain logicielle

Faire émerger des capacités de résilience cyber pour les IoT :
- Maitriser et automatiser la capacité de restauration et de mise à jour sécurisé
- Développer une capacité de détection locale

Typologie d'acteurs :
- Start-up et PME innovantes
- Laboratoires de recherche
- Partenaires industriels, utilisateurs finaux

Calendrier :
- Lancement des projets de septembre 2020 à avril 2021
- Tranche de financement spécifique jusqu'en février 2022
- Tranche de financement concurrentiel en 2022

c. *Axe vertical 3 :* la protection des petites structures contre la cybercriminalité

Les petites structures souffrent de deux problèmes principaux en cyber-sécurité : le manque de sensibilisation et l'impossibilité d'internaliser la compétence en cyber-sécurité. Par « petites structures », on peut entendre des entreprises (PME et TPE), des associations, certaines collectivités locales et même le citoyen pris individuellement.

Le manque de sensibilisation implique deux soucis majeurs. Tout d'abord, il empêche de mettre en place les méthodes simples et accessibles à tous d' « hygiène informatique ». Ensuite, il freine fortement le développement d'offres de cyber-sécurité pour ce type d'acteur (la demande étant encore relativement faible par rapport au nombre potentiel de clients – nombre qui en fait un segment de « mass market »).

L'impossibilité d'internaliser la compétence en cyber-sécurité s'explique par la petite taille de ces structures (principalement) et (à la marge pour cette taille d'acteur) par la rareté de la ressource humaine compétente sur le sujet et le prix de cette ressource. Cette impossibilité implique que le déploiement, l'exploitation et le maintien de solutions de cyber-sécurité pour ces acteurs doivent soit être opérés par un tiers (avec une difficulté à passer à l'échelle due pour le coup à la pénurie de RH), soit être totalement automatiques et transparents (ce qui n'existe pas ou très peu actuellement, en raison de plusieurs verrous technologiques). Bien entendu, à long terme, la partie « sensibilisation » à vocation à s'étendre vers de la formation pour réduire la « fracture cyber », à l'instar de ce qui s'opère pour la « fracture numérique ». En attendant, le mot clé pour les solutions de cyber-sécurité à destination de cette typologie de client est « transparence », grâce à l'automatisation du déploiement

(analyse de risque, configuration), de l'opération (protection et détection) et du maintien (mise à jour), voire de la réaction en cas d'incident.

Un troisième problème, qui a vocation à disparaitre avec les deux précédents, est le prix des solutions de cyber-sécurité, aujourd'hui très élevé. Néanmoins, ce prix (pour les petites structures) devrait baisser largement avec la consolidation de ce « mass market » (grâce à l'amortissement des coûts fixes, le déploiement simplifié grâce au SaaS, potentiellement à la valorisation des données issues de ce « mass market ») et avec l'automatisation (levant une grosse partie de la charge RH).

Ces problèmes majeurs rendent les petites structures particulièrement démunies face à la menace cyber. Elles sont ainsi des cibles de choix des cybercriminels, qui représentent la quasi-intégralité de la menace pour ces acteurs. Ainsi, notre société est divisée en trois catégories de victimes cyber potentielles : celle qui a la compétence pour se protéger, celle qui a l'argent pour payer ceux qui ont la compétence pour les protéger et celle qui est laissée dépourvue. La numérisation du monde étant exponentielle et bientôt totale, la situation ne devrait pas s'arranger d'elle-même. Cette situation n'est pas sans rappeler la « Far West » américain au moment de la « conquête de l'ouest ». Nous pourrons d'ailleurs tirer quelques leçons de cette période plus tard.

De l'autre côté du miroir se trouvent les cybercriminels qui, similairement (voire conjointement) à la criminalité dans le « monde physique », sont plus ou moins organisés de l'individu au grand groupe, en passant par de petites et moyenne entités. Il s'agit d'un acteur économique qui semble avant tout chercher la rentabilité et qui, contrairement au « monde physique », bénéficie d'une certaine impunité due au caractère transnational des réseaux. Néanmoins, cette recherche de la rentabilité (maximale) avant tout représente un atout à plusieurs niveaux pour le défenseur. Tout d'abord la stratégie d' « active cyberdefence » britannique (à ne pas confondre avec la désignation commune de d'active cyberdefense qui peut désigner des activités de « hackback ») semble très pertinente à court terme. Cette démarche, principalement orientée vers leurs équivalents OIV et OSE, consiste simplement à fournir des éléments simples et robustes de cyber-sécurité (sécurisation des mails, des sites web, etc.). En effet, le marché (pour le criminel) des cibles sans défense est très vaste aujourd'hui et n'est pas (encore) saturé par la cybercriminalité (il le serait si l'intégralité des petites structures payaient au maximum de leur capacité (financière) des rançons aux cybercriminels). Ainsi, en théorie, il suffirait qu'un segment de ce marché augmente son niveau de protection de manière raisonnable (on reste donc loin d'une cible inattaquable) pour tenter de faire baisser la rentabilité du criminel d'autant et ainsi inciter ce dernier à redéployer quasi-intégralement son activité sur un segment non encore adressé et à plus forte rentabilité. Il s'agirait donc d'un effet de seuil extrêmement efficace (avec un ROI très élevé).

Plaçons-nous un instant du côté du cybercriminel pour comprendre sa démarche. Comme une entreprise classique, le cybercriminel bénéficie d'un chiffre d'affaire (rançons, arnaques au président, etc.) et subit un coût d'exploitation (pour faire tourner son business). La différence entre les deux (l'excédent de trésorerie d'exploitation = ETE) se doit d'être positive sur le long terme pour que l'entreprise perdure (et le criminel cherchera à la maximiser). Il faut aussi déduire le coût des investissements (qui ont permis de démarrer le business ou qui tendent à l'améliorer). Une fois retranché ce dernier coût à l'excédent de trésorerie d'exploitation, nous obtenons le flux de trésorerie disponible. S'il est positif, il permet de rembourser ses dettes et de faire des bénéfices. S'il est négatif, il est nécessaire de s'endetter. Sans rentrer dans le détail du marché de la dette pour les criminels, on peut imaginer qu'il est assez peu flexible et relativement cher. Il ressort de cette courte analyse trois manières simples de se débarrasser du cybercriminel. La première est celle de l' « active cyberdefence » et consiste en un investissement relativement faible dans sa sécurité pour être moins rentable que d'autres cibles sur un marché non saturé. Très efficace à court terme, cela devient de

plus en plus couteux à moyen et long terme quand toutes les victimes potentielles adoptent la même stratégie. La deuxième approche est plus systémique et consiste à rendre le modèle économique du cybercriminel non viable (un ETE négatif), en faisant globalement augmenter le niveau de sécurité d'une grande partie des victimes potentielles. A moyen terme, cela impose au criminel des investissements pour faire à nouveau baisser ses coûts d'exploitation ou augmenter son chiffre d'affaire (par exemple en achetant sur le darkweb des malwares plus évolués). Enfin la troisième option est de faire subir au criminel un choc de trésorerie (décalant ses revenus dans le temps) le rendant temporairement incapable de maintenir son activité et de rembourser ses dettes éventuelles. Il s'agit là d'opération de grande envergure (démantèlement de botnet, indisponibilité des fournisseurs de logiciels, fermeture des « market place » pour l'acquisition de ressources, incapacité temporaire à trouver des hébergements sur le web, etc.).

Face au développement de la cyber-sécurité, tendance de long terme qui menace le modèle économique de la cybercriminalité, cette dernière s'est organisée pour faire baisser ses différents coûts et augmenter sa rentabilité. Il existe donc un véritable marché souterrain pour la cybercriminalité, où est possible l'acquisition de *malwares*, d'exploits, de frameworks d'attaque, etc. Cela permet au criminel d'opérer à la fois plus d'attaques, à plus large échelle et à moindre coût (baisse du coût d'exploitation) et d'augmenter le taux de succès grâce à des attaques plus sophistiquées (augmentation du chiffre d'affaire), moyennant l'acquisition d'un certain nombre d'outils (investissements) sur le darkweb (*market place*) par exemple. Cela favorise donc un écosystème de cybercriminels (parce qu'il s'agit bien de criminels) qui valorisent leurs compétences techniques en revendant leurs services (travail en régie ou intégration verticale dans une organisation) ou leurs outils à des criminels qui mènent les attaques. Une professionnalisation de l'ingénierie (sophistication des *malwares* et des attaques) et de l'opération (attaque à grande échelle et bien organisée) est donc observable et devrait s'accroitre à court et moyen terme. Il s'agit là d'un véritable marché secondaire de la cybercriminalité qui, dans l'absolu, n'est pas à négliger.

Objectifs :
Emergence d'une offre de cyber-sécurité pour les petits acteurs :
- *Transparente au déploiement et à l'usage*
- *Contre la cybercriminalité*
- *A coût abordable*

Typologie d'acteurs :
- *Start-up et PME innovantes*
- *Laboratoires de recherche*
- *Partenaires pour la diffusion large*

Calendrier :
- *Lancement des projets de septembre 2020 à avril 2021*
- *Tranche de financement spécifique jusqu'en février 2022*
- *Tranche de financement concurrentiel en 2022*

d. *Axe transverse 1 :* l'amorçage en cyber-sécurité

Le rapport de Wavestone 2019 (« Radar 2019 des startups de cyber-sécurité en France ») saluait une augmentation des levées de fonds dans le domaine des start-up de cyber-sécurité françaises (un volume quatre fois plus important qu'il y a deux ans). Il s'agit là d'un début, mais la route est encore

longue. En effet, l'augmentation s'explique aussi par l'arrivée à un certain niveau de maturité d'une vague d'acteurs créés il y a quelques années et pas seulement par une augmentation notable de la volonté des investisseurs de s'impliquer tôt dans le développement des jeunes sociétés. Il y a bien une amélioration, mais elle n'est pas aussi nette que ce qu'elle y parait au premier regard. En effet, le volume plus important correspondant « principalement » à quelques acteurs atteignant un niveau de maturité ou d'ancienneté cohérent avec des levées de l'ordre de 10M€. C'est une très bonne chose que quelques-unes de nos start-up atteignent ce niveau et c'est encore mieux de constater que le marché de l'investissement à ce niveau s'y intéresse, mais cela ne préjuge que peu de l'évolution des investissements sur les tranches inférieures (très « early » stage) que nous suivions jusque-là. Le montant quatre fois supérieur du volume de levée de fonds est donc principalement dû à l'apparition de « grosses séries A » et de séries B. Si l'on enlève ces nouveaux types de levées du calcul pour comparer avec les données des années précédentes, nous obtenons un volume relativement constant sur le très « early » stage depuis deux ans. Néanmoins, l'accès de nos start-up à ces levées un peu plus « late stage » étant une très bonne chose, il est intéressant d'essayer d'en déterminer les facteurs de succès. Regardons le top 4 des start-up ayant le plus levé en 2018-2019 d'après le radar Wavestone :

- Alsid (créée en juin 2016) : après un tour de « pre-seed » de 60k€ en juin 2017, levée rapide d'un tour de « seed » de 1,5M€ en septembre de la même année puis d'une série A de 13 M€ en avril 2019 (un peu moins de deux ans après). La société a donc trois ans quand elle opère sa série A. Trois points sont à noter (ce qui n'enlève rien au succès et au mérite de la société) :
    - Premièrement, il s'agit d'anciens de l'ANSSI qui ont bénéficié de la bienveillance de leur administration pour monter leur société dans les meilleures conditions. Il s'agit là d'un différentiant important à garder à l'esprit et certainement à reproduire à l'avenir.
    - Entre la création et la première levée conséquente, il s'est écoulé un an et trois mois. C'est déjà relativement long. Une société n'ayant pas bénéficié du premier point n'aurait peut-être pas aussi bien évolué. Il est essentiel de se positionner sur ce segment très « early stage » qui représente un problème généralisé.
    - Il s'écoule encore près de deux ans avant une série A conséquente qui laisse espérer un développement important (incluant l'international). C'est à la fois un bon et un mauvais signe. On y voit un signe positif : l'entreprise n'est pas dans une course à la valorisation et construit une technologie et un business solide. Néanmoins, cela traduit mécaniquement un développement (RH en particulier) relativement lent.
- Sqreen : cette startup commence avec un « pre-seed » important et intègre ensuite un des incubateurs les plus connus au monde, Y Combinator à San Francisco. Cela n'a rien de dénigrant mais malgré ses fondateurs français, il s'agit maintenant d'une start-up américaine.
- Cybelangel (créée en février 2013) : « pre-seed » de 1M€ en juillet 2015, « seed » de 3M€ en juillet 2017 puis série A de 10M€ en octobre 2018 (et ensuite une série B de 30M€ en février 2020). On retrouve l'absence de levée pendant un temps important au début (deux ans) et un temps important avant la levée suivante (deux ans encore) avec les mêmes conclusions que précédemment.
- Sentryo (créée en juillet 2014) : « seed » de 2M€ en mars 2016 (presque deux ans après la création) puis série A de 10M€ en décembre 2018 (deux ans après). Les conclusions sont les mêmes mais la série A comporte de l'ordre de 7 investisseurs (ce qui semble relativement important pour 10M€) ce qui peut sous-entendre une difficulté à lever à ce stade (et à cette date). La société a, depuis, été rachetée par CISCO, ce qui montre à la fois une confirmation de son potentiel et un éventuel manque de financement côté européen.

Le palmarès (des levées avoisinant les 10M€) se poursuit avec Odaseva (centrée sur les données en général dont un aspect est la cyber), Bleckwen (plutôt orientée fraude que cyber-sécurité) et Reachfive qui intègre une composante sécurité dans une stratégie plus globale.

Sur le nombre restreint de start-ups françaises spécialisées en cyber-sécurité, le rapport de Wavestone révélait que seules 44 % se positionnent de manière disruptives (avec un produit véritablement nouveau). Le bon côté de ce chiffre est qu'il est en augmentation, mais il en reste néanmoins que plus de la moitié des nouvelles start-up qui nous intéressent sont peu ou pas disruptives. Un certain nombre de start-up basent une partie (trop ?!) importante de leur stratégie technologique et commerciale sur l'aspect franco-français de souveraineté. Cet aspect, qui a certes vocation à être un avantage majeur pour adresser le marché intérieur, se retrouve donc être l'argument principal de vente. Cela a de très nombreux impacts négatifs. Le développement se trouve ainsi intrinsèquement limité au marché intérieur, réduisant ainsi *a priori* l'ambition commerciale de la société. De plus, cette démarche n'attire pas particulièrement l'innovation ni une approche concurrentielle basée sur une ambition technologique importante et disruptive.

Les compétences en France dans le domaine de la cyber-sécurité sont indéniablement importantes et pointues. Il est donc impératif que notre tissu de start-up spécialisées dans le domaine, ambitieuses et disruptives, le reflète. Il est même anormal que ce ne soit pas le cas. En résumé, nous avons des startups françaises ambitieuses et disruptives en cyber-sécurité mais trop peu par rapport au potentiel observable. Un certain nombre de freins identifiés par Wavestone peuvent l'expliquer. Nos experts cyber sont plus des ingénieurs pointus et averses au risque que des « serial entrepreneurs ». Cela induit deux biais. Tout d'abord une difficulté à « sauter le pas » pour créer une nouvelle entreprise et ensuite un biais technologique au détriment du commercial et du marketing. Il faut, à cela, rajouter la difficulté (identifiée plus haut) d'effectuer sa première levée de fond. La phase de l'idée (voire avant) jusqu'au premier client est une phase présentant un risque inhérent élevé dans tous les domaines technologiques. Le domaine de la cyber-sécurité étant très technique et en évolution permanente et rapide, ce risque est encore accru. Il est donc naturel pour les investisseurs de limiter ce risque en attendant qu'une société ait son premier client, garant de l'intérêt commercial et technologique de la solution. Il n'en demeure pas moins un trou de financement dans cette période parfois très longue. En particulier, comme c'est indiqué dans le rapport Wavestone, le temps moyen entre un POC (« Proof Of Concept ») réussi avec une entreprise et un contrat est de plus de 6 mois, alors que ces POC sont déjà durs et longs à obtenir. Prenons à l'opposé l'exemple de l'Etat d'Israël (la « startup nation » par excellence). Plusieurs facteurs réduisent considérablement ce trou. D'abord, les entreprises se tournent plus naturellement vers les start-up pour couvrir leurs nouveaux besoins. Ensuite, les investisseurs investissent plus tôt (et plus). Enfin, l'Etat se positionne relativement souvent sur ce créneau très « early stage ».

A titre de comparaison avec de nombreuses startups américaines ou israéliennes, on observe souvent un court délai entre la création des sociétés et la première levée, ainsi qu'un montant relativement important (en moyenne supérieur à $2M). De surcroît, l'enchainement rapide avec les levées suivantes permet une rapidité de croissance et d'exécution plus importante. Cette tendance pourrait s'accentuer à court et moyen terme avec l'explosion du marché cyber et sa maturation progressive.

L'idée est donc de mettre en place une démarche réduisant le risque (au moins perçu) des investisseurs, attirant les entreprises clientes et favorisant la création d'entreprises dans ce domaine très spécifique qu'est la cyber-sécurité. Il s'agirait donc de soutenir un projet de création d'un incubateur très « early stage » (un **startup studio** en réalité) qui devra spécifiquement :

- **Encourager les entrepreneurs potentiels à « sauter le pas » et à créer leur entreprise**

- Chercher des applications dans un domaine innovant et disruptif et impliquer la recherche académique
- Apporter des compléments de compétences commerciales et marketing pour supporter une ambition technologique par une ambition commerciale importante
- Aider à l'adaptation de l'idée aux besoins concrets du marché
- **Aider aux démarches d'obtention rapide d'un premier client par un réseau et un cadre encourageant pour les clients potentiels**
- Une aide au recrutement de talents pour favoriser un rythme soutenu de développement
- **Fournir un capital d'amorçage conséquent très tôt**
- Assurer la bienveillance et le soutien des différentes administrations dans le domaine
- Favoriser, en plus des aspects technologiques, les aspects d'expérience utilisateur, essentiels au marketing

Objectifs :
Soutien à la création d'un « startup studio » co-localisé entre Rennes et Paris pour :
- Stimuler la création de startups en cyber-sécurité en capitalisant sur la co-localisation avec la cyberdefense factory et le campus cyber
- Réduire le temps d'acquisition du premier client

Calendrier :
- Lancement des antennes de Rennes et Paris d'ici mi 2021

e. *Axe Transverse 2 :* la problématique des données cyber

Le sujet des données en cyber-sécurité intègre trois problématiques complémentaires au travers de trois typologies d'acteurs : les producteurs de données d'intérêt cyber (industriels, utilisateurs, individus, etc.), les exploitants de données cyber (les industriels de la cyber-sécurité, dans une certaine mesure, l'ANSSI et le CALID, etc.) et la R&D nécessitant des données. Dans tous les cas, il s'agit bien de volumes extrêmement importants (i.e. « big data »), représentant une difficulté de captation, de stockage et de traitement. De surcroit, les aspects liés à la vie privée et au RGPD encadrent un certain nombre de ces données (celles à caractère personnel ou associées) mais ce n'est pas le cas de toutes (on peut penser aux logs de systèmes industriels par exemple).

La première catégorie d'acteurs est principalement impactée par deux aspects : la maitrise des données (de leur propriété, de leur diffusion, etc.) et leur valorisation (technique pour en extraire des informations et éventuellement financière). Trois situations sous optimales apparaissent fréquemment, reflétant partiellement le niveau de maturité en cyber-sécurité. Premièrement, les données peuvent ne pas être stockées (par manque de sensibilisation à leur valeur intrinsèque, par contrainte budgétaire, etc.). Les données sont perdues et leur valorisation est au plus bas. Deuxième possibilité, elles sont stockées mais non exploitées. Bien souvent, les problématiques d'espace de stockage et de contraintes réglementaires impliquent une « rotation » des données et ne fait que décaler dans le temps leur disparition. Le potentiel de valorisation est alors plus haut (en raison de la disponibilité) mais le résultat reste le même sur la durée. La troisième éventualité permet une valorisation souvent partielle. Il s'agit du cas où le traitement des données est confié à un tiers. Elles sont captées, stockées et traitées par ce tiers. Cela peut présenter deux problèmes. En fonction des capacités du tiers, la valorisation peut être limitée (i.e. l'intégralité des informations latentes n'est pas extraite). De plus, certains tiers, en particulier non nationaux, ne proposent aucunes garanties sur la

maitrise des données ce qui aboutit souvent à la perte de propriété. En particulier au niveau de nos acteurs français, la perte de souveraineté de ces données correspond à terme à une perte économique (ne serait-ce que pour leur valeur financière) majeure. Aujourd'hui, notre « balance informationnelle » sur les données cyber est très négative. Cette balance implique une dépendance (comme peut l'impliquer la balance commerciale) sur cette matière première essentielle à nos capacités de cyber-sécurité. La crise actuelle du Covid19 a pu mettre en valeur comment certaines dépendances peuvent poser de gros soucis.

La deuxième catégorie d'acteurs (les exploitants de données) est confrontée à des difficultés de représentativité (données spécifiques à leurs clients, volumes insuffisants, etc.), à des besoins de modèles et d'algorithmes innovants pour le traitement et, parfois (dans les cas que nous développerons), à des contraintes de maitrise des données fournissant des garanties aux clients. Les difficultés de représentativité sont critiques en particulier d'un point de vue compétitivité commerciale. En effet, cet aspect induit un niveau de pertinence de la prestation et des modèles (d'apprentissage par exemple). Pour simplifier, si, par exemple, les algorithmes d'apprentissage machine sont relativement connus et que leur implémentation et application tendent à se standardiser, la différence entre les acteurs se fera sur la base d'apprentissage. La capacité d'avoir accès à une base représentative tant en volume qu'en diversité de données devient donc un élément de souveraineté à part entière. Or, aujourd'hui, nos acteurs cyber dépendent tous de leur base interne (potentiellement biaisée par la spécificité de leurs clients) loin d'atteindre une taille critique. Face à des acteurs qui ont naturellement accès à beaucoup de données diverses (par exemple CISCO, Google, Microsoft, Apple, etc.) cette démarche est trop risquée (pour ne pas dire vouée à l'échec). Il est clair qu'une approche nationale, et européenne à terme, serait nécessaire pour rester compétitif face à ces acteurs et ainsi conserver nos capacités propres de cyber-sécurité.

La troisième catégorie d'acteurs (la R&D et l'innovation) permet d'enrichir les capacités de traitement et d'exploitation des données et plus généralement d'apporter de nouvelles solutions de cyber-sécurité. L'enjeu est donc, bien sûr, de rester pertinent et compétitif. Néanmoins, aujourd'hui, très peu de projets innovants (aucuns, en fait, n'ont été observés dans le cadre du Grand Défi) ne nécessitent pas d'accès à des données représentatives pour tester, entrainer ou développer les théories, les modèles et les produits. Il s'agit là d'un élément essentiel pour passer d'une idée à un produit opérationnel et commercialisable. Avoir un accès à une base de données variée et représentative donnerait un avantage important à tous nos acteurs innovants, start-up, PME, grands groupes, mais aussi laboratoires de recherche et organismes de transfert de technologie.

Il en ressort que la mise en place, au niveau national voire, à terme, européen, d'une base de données garantissant à la fois la maitrise et la valorisation des données à ceux qui les produisent, l'accès à des données représentatives (avec un aspect de volume critique) à ceux qui les exploitent et une capacité d'expérimentation à ceux qui innovent, est un enjeu de souveraineté. Le Grand Défi est particulièrement impacté par le dernier aspect sur la capacité d'expérimenter pour l'innovation. Néanmoins, la question est globalement plus large. En particulier, les acteurs cyber exploitant ces données sont largement représentés au sein du CSF « industrie de sécurité » et du futur campus cyber, lieu fédérateur de la cyber-sécurité française qui serait idéal pour accueillir cette base de données. Les producteurs de données sont très variés mais il pourrait être pertinent d'intégrer en premier lieu les OIV et OSE tout en invitant toute personne ou entité intéressée (et porteuse de données) à intégrer le projet. Ce très vaste périmètre et cet enjeu de souveraineté font de ce projet un candidat idéal pour être financé dans le cadre du pacte productif et du PIA4. Sous réserve de ces financements, le Grand Défi pourrait porter le projet initial.

En complément, une piste plus prospective pouvant participer à adresser cette problématique des données serait d'investiguer un domaine plus amont : la génération de données synthétiques. Tout l'enjeu, ici, est d'arriver à obtenir la représentativité nécessaire en partant d'une modélisation et d'un échantillon restreint.

Objectifs :

Création d'une plateforme de données à vocation européenne à fin de :

- *Garantir la maitrise de la propriété des données*
- *Permettre une représentativité statistique pertinente*
- *Stimuler l'innovation*

Typologie d'acteurs :

- *Industriels propriétaires de données d'intérêt cyber*
- *Industriels spécialisés dans l'exploitation de données cyber*
- *Acteurs cyber innovants*