NOTE D'INFORMATION ET D'ACTUALITÉ DU CIGREF

RECOMMANDATIONS AU SUJET DES IA GÉNÉRATIVES

ISSUES DE NOTRE TASKFORCE IA GÉNÉRATIVES

Cigref
RÉUSSIR
LE NUMÉRIQUE

2023

Cigref

Note d'information et d'actualité

Recommandations au sujet des IA génératives

Juillet 2023

TABLE DES MATIÈRES

1 INTRODUCTION

La mise à disposition de l'application ChatGPT (*Chat Generative Pre-trained Transformer*) auprès du grand public, en novembre 2022, a créé une tempête médiatique dans le champ des technologies d'intelligence artificielle et généré un véritable engouement. En effet, ChatGPT a rendu visible la dynamique de recherche en intelligence artificielle qui se développe depuis plusieurs années déjà et dont les conséquences sont annoncées et documentées. C'est désormais une évidence, les outils d'IA générative auront des impacts importants sur les administrations publiques et les entreprises de très nombreux secteurs d'activité, avec des effets systématiques sur leur productivité et leurs performances.

C'est pourquoi le Cigref a lancé une *Task Force,* pilotée par Baladji SOUSSILANE, *Vice-President Digital & IT* du groupe Air Liquide, animée par Marine de Sury, directrice de mission au Cigref et réunissant plus d'une quarantaine de ses membres pour échanger sur leurs pratiques et mettre en commun leurs expériences. Cette Note d'Information et d'Actualité du Cigref (NIAC) a pour objectif de lister les différentes recommandations préconisées par le Cigref afin que ses membres mais aussi potentiellement d'autres organisations, entreprises, administrations, académies ou associations, etc. puissent les personnaliser en fonction de leur contexte et de leurs enjeux.

Les entreprises réagissent de diverses façons face à l'utilisation des outils d'IA générative. Certaines préfèrent interdire à ce stade, tout usage en interne employant des données de l'entreprise ou l'ouverture de comptes avec l'adresse mail professionnelle des collaborateurs, afin d'éviter l'exfiltration des données stratégiques ou sensibles. D'autres, à l'inverse, en profitent pour créer de l'appétence pour ces nouvelles technologies et générer des opportunités business. Pour cela, elles partagent en interne des outils d'IA générative privés en mode SaaS ou hébergés en leur sein, donnent des lignes de conduite indiquant ce qu'il est possible ou non de faire (par exemple, interdiction d'utiliser des documents d'entreprises sur les outils publics) et mettent en place une *"tour de contrôle"* pour réguler les usages.

Certains acteurs proposent des services surfant sur la vague « IA Générative » qui demandent de télécharger pour analyse des documents et qui n'offrent aucune garantie sur l'usage des informations contenues. Par exemple, ceux de type « conservations avec vos PDF » qui impliquent que ces fichiers soient soumis à la plateforme Cloud ; Il en est de même avec des outils de type ChatGPT lorsqu'ils entraînent les LLM (*Large Language Machine*) avec des conversations/corpus de données de l'entreprise.

Quel que soit leur positionnement, toutes les organisations sont unanimes pour dire que le plus gros risque est de passer à côté ou d'être en retard sur la transformation induite par les IA génératives. **Les risques et la sécurité sont à gérer en parallèle et non en préalable à la réflexion sur les opportunités.**

La première partie de cette NIAC liste les recommandations d'utilisations et les bonnes pratiques concernant l'usage des systèmes d'IA générative qu'il est important de partager en interne. Les IA génératives déplacent dès à présent les frontières dans la productivité et la créativité, et offrent donc de véritables opportunités à saisir. Cependant, elles présentent également des risques qu'il faut identifier afin de mieux s'en prémunir. C'est l'objet de la deuxième partie de ce document.

2 RECOMMANDATIONS D'UTILISATIONS ET DE BONNES PRATIQUES

2.1 EXPÉRIMENTER LES OUTILS D'IA GÉNÉRATIVE

Pour la majorité des entreprises et administrations publiques, il apparaît indispensable de se saisir, dans les meilleurs délais, des outils d'IA générative pour s'inscrire dans une démarche **d'apprentissage, d'expérimentation et d'appropriation** afin de savoir les implémenter avec le niveau de sécurité souhaité, et d'en développer les usages appropriés et pertinents. Ensuite, la pratique permet de mettre en place des processus de validation des cas d'usage et de mieux maîtriser les coûts associés.

Plusieurs entreprises préconisent de commencer les expérimentations à petite échelle et de favoriser l'utilisation de plusieurs "petits cas d'usage" d'IA générative plutôt qu'un seul grand modèle à tout faire, en raison notamment de l'agilité, la modularité et la rapidité avec laquelle les solutions évoluent. Ce serait dommage d'investir trop tôt dans une technologie ou dans un développement interne qui serait dépassé quelques semaines plus tard.

Tester les systèmes d'IA générative contribue à **identifier leur potentiel mais aussi leurs limites**. Afin d'apprendre à les manipuler avec précaution, des organisations s'appuient sur des hommes et des femmes pour vérifier ou participer à la validation des résultats, à chaque fois que des décisions importantes sont prises sur la base de système d'IA.

2.2 ACCULTURER LES COLLABORATEURS

Au préalable, les organisations préconisent d'**évaluer la maturité** de l'ensemble des équipes sur la data et les IA afin d'adapter au mieux la communication et la formation à mettre en œuvre. En effet, les entreprises et administrations publiques sont unanimes sur l'importance de sensibiliser, acculturer et former, aux usages des outils d'IA générative. Le comité exécutif est la première cible et beaucoup d'organisations les informent régulièrement des avancées sur les différents travaux autour de ces outils.

2.2.1 DÉFINIR LA STRATÉGIE DE COMMUNICATION

L'engouement du grand public pour les systèmes d'IA générative les plus connus, ChatGPT, DALL-E, Midjourney et GitHub Copilot conduit les entreprises et administrations publiques à acculturer tous les salariés sur l'ensemble de ces outils et systèmes, en informant, démystifiant, et suscitant l'usage, ainsi qu'à partager largement des règles de conduite et d'éthique, parfois appelées les *"do's and don'ts"* sur les canaux de communication adéquats et avec des formats adaptés à la cible. Une entreprise a par exemple mis en place une alerte dédiée aux sites de type IA, via le proxy d'entreprise qui rappelle les règles dès qu'un collaborateur s'apprête à utiliser un de ces outils.

De la même façon, il serait intéressant d'imposer rapidement des **règles spécifiques aux sous-traitants**, notamment afin de garantir qu'eux-mêmes n'utilisent pas des IA publiques, par exemple sur le développement.

Plusieurs organisations présentent les concepts d'IA générative aux différentes strates de l'organisation (CxO, métiers, support, contrôles) en adaptant les messages, enjeux et objectifs à la cible.

Au regard de l'importance des corpus de données utilisés pour entraîner les IA, les entreprises s'organisent **pour assurer leur qualité et leur mise en visibilité à l'échelle**.

2.2.2 FORMER

Plusieurs organisations ont mis en place des formations sur les systèmes d'IA générative, sur les nouvelles tâches/missions induites (par exemple *prompt engineering*[1]), les risques, etc., ou bien elles orientent leurs salariés vers des ressources en ligne pour les sensibiliser.

Les équipes qui lancent des projets data et IA doivent être certaines de la qualité de l'origine, et de la fiabilité des données, et comprendre la sémantique et la gouvernance mise en place. Tous les niveaux sont impactés. La collecte des données se fait souvent au niveau opérationnel, d'où l'importance de bien faire comprendre les enjeux autour de celles-ci (qualité, complétude, …) en particulier auprès des opérationnels. Il faut que les modèles d'IA soient éthiques et explicables. Enfin, le produit généré doit apporter de la valeur et répondre aux besoins et à la stratégie de l'entreprise / administration publique.

En amont, les formations sont également à intégrer dans les formations académiques et de reconversion.

2.2.3 ORGANISER

Quelques entreprises ont mis en place un **comité transverse et pluridisciplinaire** aussi appelé "*AI Tower*", ou encore « *Generative AI committee* », pour suivre et valider les différents cas d'usage d'IA génératives et LLM dans l'entreprise, comme cela a été fait pour la data, et maintenir à jour les recommandations et réponses aux *FAQ (*questions fréquemment posées*)*. Ce comité pluridisciplinaire regroupe souvent des collaborateurs des équipes cybersécurité, juridique, produits, systèmes d'information. Il constitue une bonne pratique pour faire **gagner en compétence collectivement** et pour assurer une bonne visibilité des actions en interne. Il assure également la **mise à jour des règles de conduite**. Il importe en effet de valider que le choix des fournisseurs et des modèles sont adaptés aux usages (localisation géographique, immunité aux lois extraterritoriales), que les licences permettent un usage commercial, et enfin, que les niveaux de service sont compatibles avec les exigences, bref de mesurer les impacts avant la mise en production.

[1] Le Prompt Engineering est le processus de conception et de création de prompts, ou de données d'entrée, pour conduire l'IA à effectuer une tâche spécifique. Ceci implique de sélectionner le type de données adéquat et de le formater pour que le modèle le comprenne et l'utilise. L'objectif est de créer des données de haute qualité pour permettre à l'IA d'effectuer des prédictions précises et de prendre les bonnes décisions.

Une entreprise a par exemple mis en place une *task force* pluridisciplinaire, missionnée pour faire de la veille technologique, travailler sur les cas d'usage, le démonstrateur, les risques et les aspects juridiques ainsi que sur la sensibilisation et la formation. Ainsi la maturité progresse collectivement en incluant différents métiers qui essaiment chacun de leur côté.

2.2.4 IMPLIQUER LES RESSOURCES HUMAINES

Plusieurs entreprises ont impliqué les ressources humaines afin qu'elles **identifient les nouveaux rôles et missions** émergeant avec les cas d'usage des IA génératives (superviseurs de bots et de décisions d'IA, *prompt engineer*, etc.).

Les ressources humaines cherchent également à **identifier les métiers qui vont être impactés** par l'IA afin d'anticiper les adaptations à faire ou préparer les reconversions. En effet, si le message des acteurs de ses technologies se veut rassurant et met en avant un rôle de super assistant (ou copilote) pour donner des « super pouvoirs » aux collaborateurs, les gains en efficacité promis par ces technologies requièrent d'anticiper les conséquences possibles sur une masse salariale assez peu touchée jusqu'à présent.

2.3 IDENTIFIER LES CAS D'USAGE ET LES AMÉLIORATIONS DE PERFORMANCE

Tous les cabinets de la place cherchent à vendre « leur démarche » d'identification des bons cas d'usage mais les acteurs restent prudents sur le recours à ces cabinets. Certaines organisations préfèrent mener seules la recherche des « cas d'usage en or » qui pourraient leur donner un avantage concurrentiel majeur. C'est pourquoi, ces dernières choisissent soit une approche *top down* soit une approche *bottom up*.

Approche *top down* : Certaines entreprises et administrations publiques organisent des réflexions stratégiques sur la chaîne de valeur et les usages possibles de l'IA générative pour susciter des cas d'usages avec des systèmes d'IA générative. L'idée est d'y travailler dans les métiers, sur les offres clients, les modèles d'affaires.

Approche *bottom up* : d'autres organisations font remonter les idées des collaborateurs par le biais d'expérimentations. Dans un deuxième temps, elles cherchent à identifier ceux à forte valeur ajoutée. Pour cela, certaines d'entre elles montent une *task force* pluridisciplinaire (équipes juridique, *AI factory*, DPO) qui distingue les cas d'usage prometteurs puis les analyse et enfin les teste avec les parties prenantes concernées. Le partage de REX et d'expérimentations en cours est une bonne façon de stimuler la sérendipité auprès des équipes.

2.4 METTRE EN PLACE DES PROCESSUS POUR INDUSTRIALISER DES CAS D'USAGE AVEC LES SYSTÈMES D'IA GÉNÉRATIVE

Les entreprises qui ont monté une *task force* pluridisciplinaire, fortes de leur suivi et de leur apprentissage sur les cas d'usage en cours d'implémentation, posent/construisent un cadre afin de délimiter l'implémentation de PoCs et d'anticiper leur bonne industrialisation en cas de succès.

Par exemple, une entreprise a demandé à ses développeurs de signaler leur volonté d'utiliser un outil d'IA générative, avant même que la première ligne de code ne soit écrite, afin d'évaluer la sécurité du projet. Une autre vérifie systématiquement la qualité et la bonne qualification des données utilisées avant son démarrage. Une autre encore cherche à déterminer des indicateurs/critères pour en évaluer les bénéfices, que ce soit d'un point de vue quantitatif ou qualitatif, dans l'objectif de les pérenniser au niveau de l'organisation. Enfin une dernière organisation a mis en place un comité éthique interne qui évalue l'éthique de chacun des projets, analyse les possibilités de biais ou de résultats erronés et définit le contour et les fonctionnalités du produit qu'on veut créer, ainsi que le cadre juridique associé.

2.5 SENSIBILISER POUR UN USAGE PRUDENT ET RAISONNÉ

La plupart des entreprises et administrations publiques se prononcent pour un usage prudent et raisonné de ces outils dès lors qu'il s'agit de manipuler des données sensibles - données financières, commerciales, stratégiques, de R&D, et secret de fabrication - afin d'en garantir la confidentialité. Cela nécessite d'informer les collaborateurs des risques potentiels en fonction de la classification des données. De la même manière, une certaine prudence doit s'exercer en matière d'exploitation des résultats produits par ces outils d'IA générative, afin de se prémunir des erreurs d'analyse et de traitement qui ne sont manifestement pas si rares. Il est essentiel, par ailleurs, de vérifier la conformité au RGPD du traitement des données à caractère personnel lorsque ce type d'outil est utilisé à cet effet. Enfin, il semble nécessaire de porter une attention particulière et renforcée aux enjeux de propriété intellectuelle et de droit d'auteur, dans la mesure où certaines de ces intelligences artificielles peuvent avoir été entraînées sur des corpus extrêmement vastes sans garantie robuste en la matière.

2.6 INTÉGRER DÈS LE DÉBUT L'ÉTUDE DE L'IMPACT ENVIRONNEMENTAL

Les organisations cherchent à minimiser l'impact environnemental et donc à mesurer l'impact RSE (Responsabilité Sociétale des Entreprises) de la mise en place d'une IA. Cependant, faute de maturité collective sur le sujet, elles sont plutôt démunies même si elles tentent de contrôler et de maîtriser l'utilisation des ressources (stockage, calcul par exemple). En effet, les fournisseurs, à ce stade, se contentent de réponses liées à leur politique environnementale globale.

2.7 FAIRE UN ÉTAT DES LIEUX DES SOLUTIONS ET DES FOURNISSEURS

Tout d'abord, les organisations cherchent à identifier les différents acteurs, tester les solutions dans cet environnement foisonnant et évoluant rapidement afin de les comparer pour un *benchmark*. Elles étudient le risque de dépendance des systèmes d'IA générative et les conditions générales de vente associées, afin de les adapter à leurs besoins, si nécessaire. La sensibilité des données constitue un critère important dans le choix du fournisseur. Les entreprises identifient les modèles existants pertinents pour les cas d'usage et les adaptent avec leurs données pour leurs propres cas. Par exemple, des organisations ont choisi d'utiliser les systèmes de LLM sur une instance de l'entreprise sans lien avec la version publique. Quelques entreprises choisissent des fournisseurs qui acceptent de travailler en synergie avec leur écosystème.

Les solutions *open source*, qui progressent très vite, font bien sûr partie des solutions évaluées. Elles sont identifiées comme des solutions qui ne captent ni la propriété des données des clients, ni celle des résultats. Cependant, les entreprises analysent auparavant la licence associée, et vérifient qui la détient. Elles s'assurent également qu'elles peuvent utiliser ces solutions pour une activité commerciale.

D'autres regardent dans quelle mesure les outils d'IA générative sont soumis à des lois à portée extra européenne et donc potentiellement vulnérables à des activités d'intelligence économique. Un modèle *open source* hébergé en interne est un recours possible pour des cas d'usage plus sensibles. Sur ce sujet, le CSF « Numérique de confiance » a pour objectif de faire maturer des solutions européennes afin d'assurer un vrai choix aux utilisateurs.

Enfin, certaines organisations cherchent à déterminer les fournisseurs de solutions qui utilisent eux-mêmes des outils d'IA générative et à analyser comment ces outils-là sont utilisés.

3 RISQUES LIÉS AUX SYSTEMES D'IA GÉNÉRATIVE

Lors de cet atelier, nous avons cherché à identifier les risques et les avons listés dans le schéma ci-dessous.

Risques liés aux systèmes d'IA générative

À PROPOS DU CIGREF

Au service de la croissance économique et de la compétitivité de nos membres, grandes entreprises et administrations publiques françaises, utilisatrices de solutions et services numériques, par la réussite du numérique.
Le Cigref est un réseau de grandes entreprises et administrations publiques françaises qui a pour mission de développer la capacité de ses membres à intégrer et maîtriser le numérique. Par la qualité de sa réflexion et la représentativité de ses membres, il est un acteur fédérateur de la société numérique. Association loi 1901 créée en 1970, le Cigref n'exerce aucune activité lucrative.

Pour réussir sa mission, le Cigref s'appuie sur trois métiers, qui font sa singularité.

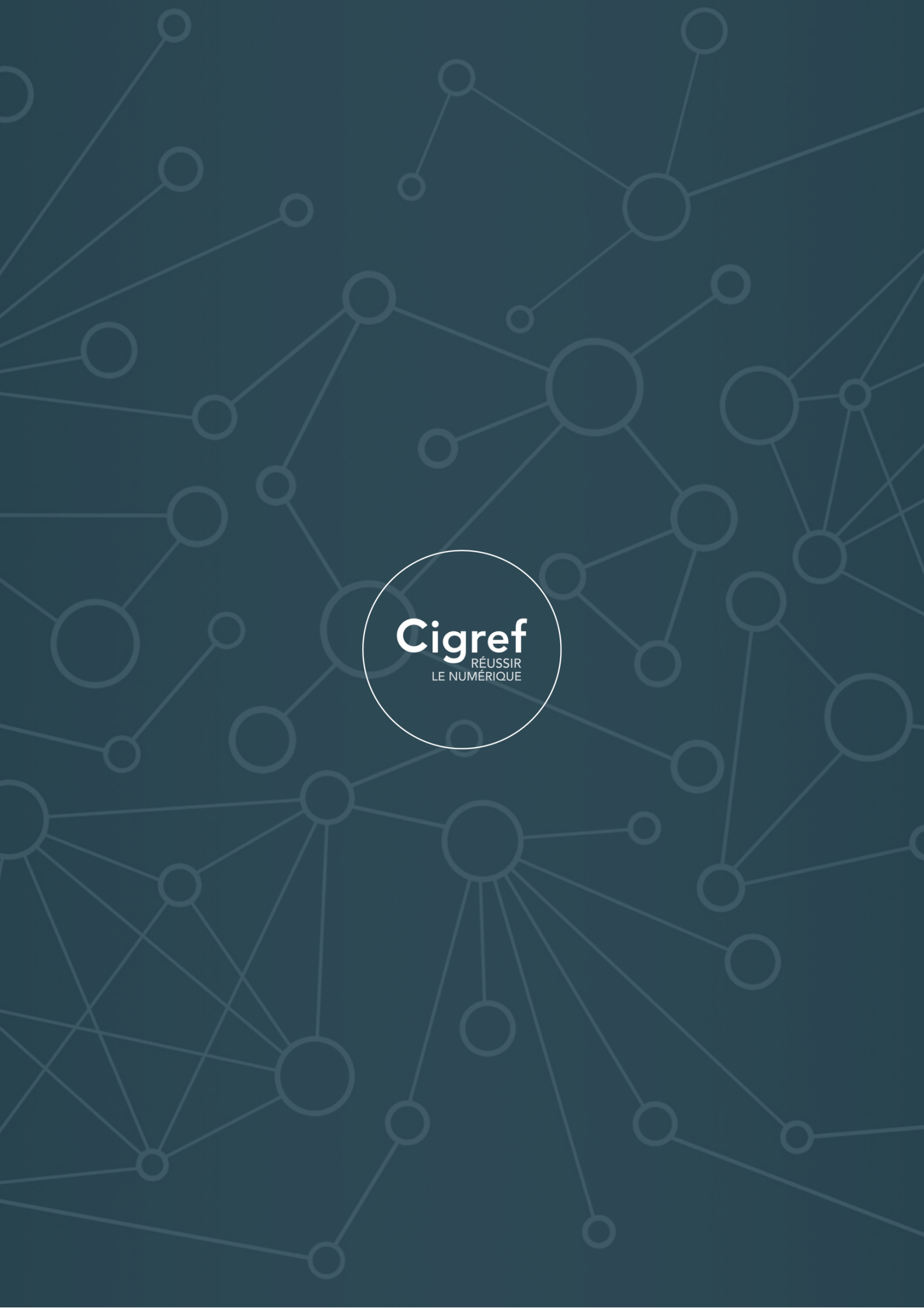
Appartenance
Le Cigref incarne une parole collective des grandes entreprises et administrations françaises autour du numérique. Ses membres partagent leurs expériences de l'utilisation des technologies au sein de groupes de travail afin de faire émerger les meilleures pratiques.

Intelligence
Le Cigref participe aux réflexions collectives sur les enjeux économiques et sociétaux des technologies de l'information. Fondé il y a près de 50 ans, étant l'une des plus anciennes associations numériques en France, il tire sa légitimité à la fois de son histoire et de sa maîtrise des sujets techniques, socle de compétences de savoir-faire, fondements du numérique.

Influence
Le Cigref fait connaître et respecter les intérêts légitimes de ses entreprises membres. Instance indépendante d'échange et de production entre praticiens et acteurs, il est une référence reconnue par tout son écosystème.

NOUS CONTACTER

www.Cigref.fr
21 av. de Messine, 75008 Paris
+33 1 56 59 70 00
Cigref@Cigref.fr


Cigref
RÉUSSIR
LE NUMÉRIQUE