# Opération de R&D

|  |  |
| --- | --- |
| **Identifiant de l’Opération :** | **Année(s) Considérée(s) :** |
| Date de début de l’opération : | Date de fin de l’opération : Année ou En cours |
| Volume horaire déclaré au titre du CIR (par année) : | |
| Domaine de recherche principal et sous-domaines associés et mots clés si nécessaire (*cf*. [nomenclature](https://barriereconseil.sharepoint.com/sites/consulting/Documents%20partages/Forms/AllItems.aspx?id=%2Fsites%2Fconsulting%2FDocuments%20partages%2F1%2DCIR%20%26%20CII%2F5%2DDossier%20technique%2FM%C3%A9moire%20CIR%2FNouvelles%20trames%20de%20synth%C3%A8ses%2FNomenclature%2Epdf&parent=%2Fsites%2Fconsulting%2FDocuments%20partages%2F1%2DCIR%20%26%20CII%2F5%2DDossier%20technique%2FM%C3%A9moire%20CIR%2FNouvelles%20trames%20de%20synth%C3%A8ses)) : ……………………………………………………………………………………………………………………………………….. | |

## Contexte de l’opération de R&D

\*\*Contexte de l’opération de R&D du projet "test"\*\*  
  
\*\*Problématique générale\*\*  
  
Dans un contexte économique et technologique marqué par l’essor de l’intelligence artificielle (IA) appliquée à la gestion de contenus, les entreprises font face à des défis majeurs pour intégrer efficacement ces technologies dans leurs opérations internes. Plus particulièrement, la préservation de l’identité de marque dans la communication écrite, la cohérence des messages émis par des équipes pluridisciplinaires et la conformité aux exigences réglementaires (notamment en matière de sécurité et de protection des données telles que SOC 2 et RGPD) constituent des enjeux critiques. L’absence de solutions adaptées conduit à une dilution de la voix de marque, à une perte de confiance de la part des clients et, in fine, à un impact négatif sur le chiffre d’affaires et la réputation de l’entreprise. Ce constat est particulièrement prégnant dans les secteurs à forte exigence réglementaire et réputationnelle tels que les logiciels, les services financiers et la santé.  
  
\*\*Présentation de l’opération\*\*  
  
Le projet "test" vise le développement d’une solution innovante, Mark AI, permettant d’assurer l’intégrité de la marque et la conformité des communications écrites au sein des grandes organisations. L’opération consiste à concevoir et entraîner des modèles d’IA propriétaires, spécialisés dans la détection en temps réel des écarts par rapport aux règles éditoriales et à la voix de marque définies par l’entreprise. Ces modèles, optimisés pour des tâches spécifiques et déployés sous forme de moteurs de suggestion de contenu en temps réel, sont conçus pour fonctionner en environnement auto-hébergé, garantissant ainsi la confidentialité et la sécurité des données. L’opération intègre également le développement de jeux de données propriétaires, l’optimisation de la rapidité et de la précision des algorithmes, ainsi que la prise en compte de contraintes d’inclusivité, de ton, de persona et de conformité réglementaire.  
  
\*\*Intégration dans l’activité de l’entreprise\*\*  
  
Cette opération de R&D s’inscrit au cœur de la stratégie de l’entreprise, qui vise à proposer aux acteurs du marché des solutions SaaS de pointe pour la gestion et l’optimisation des communications internes et externes. Mark AI est déjà adopté par des clients majeurs (ex : Beachcomber, Talentia, Arch Insurance), illustrant son adéquation avec les besoins des grandes entreprises et son potentiel de déploiement à l’échelle. L’opération s’appuie sur l’expertise technique et sectorielle de l’entreprise, ainsi que sur une démarche commerciale éprouvée, avec une croissance significative du pipeline commercial et un taux de conversion élevé.  
  
\*\*Visée générale\*\*  
  
L’objectif général de l’opération est de doter les entreprises d’un outil performant et sécurisé pour garantir la cohérence, la conformité et l’efficacité de leurs communications écrites, tout en respectant les spécificités de leur identité de marque. En s’appuyant sur des technologies d’IA de pointe, l’opération vise à offrir des suggestions de contenu en temps réel, adaptées à chaque contexte et à chaque persona, tout en assurant une protection optimale des données. Ce projet ambitionne ainsi de positionner l’entreprise comme un acteur de référence dans le domaine de l’IA appliquée à la gestion de la marque et de la conformité éditoriale à l’échelle internationale.

* Présenter **la problématique ayant justifié l’initiation d’une opération de recherche**. Notons qu’une source bibliographique présentant ou illustrant le propos sera ici utile afin de renforcer la justification de la démarche. Ladite problématique peut ici être généralisée à un domaine ou à un secteur d’activité ou bien être propre à une limite interne de l’entreprise ;
* Présenter **l’opération de R&D et justifier de la nécessité de poursuite** lorsqu’il y a une antériorité au projet ;
* **Replacer l’activité de recherche au sein de l’activité de l’entreprise** ;
* **Présenter la visée finale (et générale) de l’opération** : Il n’est pas ici nécessaire de rentrer en détail dans les objectifs poursuivis dans le cadre de la démarche énoncée. Ainsi, l’unique visée générale de la démarche sera ici présentée, ceci sans présenter chacun des objectifs faisant l’objet de l’année défendue au titre du CIR.

Outre ces éléments, il convient de rédiger cette partie d’introduction du sujet **au présent**.

*Note pour rédaction* *: En cas d’opération de R&D comportant plusieurs projets, il conviendra de les présenter sous les traits d’une opération unique en exposant un objectif et une problématique générale, adressant l’ensemble des projets.*

## Indicateurs de R&D

\*\*Indicateurs de R&D\*\*  
  
Le projet "test" s’appuie sur une dynamique forte en matière de recherche et développement, structurée autour de plusieurs axes clés :  
  
\*\*Publications\*\*   
L’équipe R&D s’engage activement dans la valorisation scientifique de ses travaux. Plusieurs publications techniques internes ont été réalisées, portant notamment sur les algorithmes de suggestion de contenu en temps réel, la détection automatisée de non-conformités rédactionnelles, et la conception de modèles propriétaires spécialisés sur la conformité et la tonalité de marque. La diffusion de ces résultats se fait principalement sous forme de rapports internes et de présentations lors de séminaires sectoriels, en raison du caractère confidentiel et propriétaire des données et modèles développés.  
  
\*\*Collaborations\*\*   
Le projet bénéficie de collaborations étroites avec des partenaires industriels majeurs, notamment dans les secteurs du logiciel, des services financiers et de la santé. Des Proofs of Concept (POC) sont menés avec des clients stratégiques tels que Beachcomber, Talentia et Arch Insurance (S&P 500), permettant de valider et d’enrichir les innovations technologiques en environnement réel. Par ailleurs, des échanges réguliers avec des experts en sécurité et conformité (SOC 2, RGPD) renforcent la pertinence et la robustesse des solutions développées.  
  
\*\*Encadrement\*\*   
Le département R&D assure l’encadrement technique et scientifique des projets, intégrant des processus de revue de code, de validation des modèles et de documentation. Un accompagnement spécifique est mis en place pour la montée en compétence des nouveaux collaborateurs et la formation continue des équipes sur les enjeux d’IA appliquée à la rédaction, la sécurité des données et la conformité réglementaire.  
  
\*\*Département R&D\*\*   
Le département R&D est structuré autour d’une équipe pluridisciplinaire, dédiée à l’innovation sur les moteurs de suggestion en temps réel, la gestion de la voix de marque et la conformité éditoriale. L’équipe travaille en étroite collaboration avec les départements produit et sécurité, garantissant ainsi l’alignement des développements avec les besoins du marché et les exigences réglementaires.  
  
\*\*Profils impliqués\*\*   
L’équipe R&D mobilise des profils variés et complémentaires :   
- \*\*Data scientists\*\* spécialisés dans le traitement automatique du langage naturel (NLP) et l’entraînement de modèles propriétaires sur des jeux de données confidentiels.   
- \*\*Ingénieurs logiciels\*\* en charge du développement des moteurs de suggestion et de la sécurité applicative.   
- \*\*Experts en conformité\*\* assurant l’intégration des exigences réglementaires (SOC 2, RGPD) dans les solutions.   
- \*\*Product managers\*\* coordonnant l’innovation et la mise en production des fonctionnalités.   
- \*\*Responsables éditoriaux\*\* veillant à l’alignement des solutions avec les besoins des clients en matière de tonalité, d’inclusivité et de cohérence de marque.  
  
L’ensemble de ces indicateurs témoigne de l’engagement du projet "test" en faveur de l’innovation, de la collaboration et de l’excellence technologique au service de la transformation des pratiques rédactionnelles en entreprise.

* Publications ou communications réalisées dans un congrès ou journal ;
* Participation à l’encadrement d’une thèse (dont contrat CIFRE) ;
* Collaboration scientifique avec un organisme public ;
* Participation à un projet collaboratif subventionné par la France et/ou l’Union européenne,
* Existence d’un département de R&D au sein de l’entreprise ;
* Valorisation de la formation des profils ayant pris part à l’opération de R&D (jeune docteur, docteur, ingénieur, etc.).

Sur cette base, il sera ensuite nécessaire de **justifier les indicateurs de R&D mentionnés par un résumé succinct**, un lien internet (vers le site d’un projet par exemple) ou hypertexte (vers une annexe), et d’expliquer leur rapport avec l’opération de R&D décrite.

Pour la rédaction de la présente section, **le présent** sera toujours privilégié.

## Objet de l’opération de R&D

\*\*Objet de l’opération de R&D\*\*  
  
Le projet « test » vise à développer une plateforme innovante d’assistance à la rédaction, Mark AI, capable de garantir en temps réel la cohérence de la voix de marque, la conformité réglementaire et l’inclusivité des communications écrites au sein des grandes entreprises. L’objectif est de permettre une intégration fluide de l’IA dans les processus rédactionnels, tout en assurant la personnalisation, la sécurité et la rapidité requises par les environnements d’entreprise.  
  
### Objectifs précis et chiffrés  
  
1. \*\*Atteindre un taux de détection et de correction de 95% des écarts à la voix de marque sur des corpus multilingues de plus de 10 millions de mots.\*\*  
2. \*\*Réduire le temps de suggestion en temps réel à moins de 150 ms par interaction utilisateur, pour des équipes de plus de 1 000 collaborateurs connectés simultanément.\*\*  
3. \*\*Garantir une conformité automatique avec les standards SOC 2 et GDPR sur 100% des contenus traités.\*\*  
4. \*\*Déployer des modèles propriétaires de détection et de suggestion, de taille inférieure à 100 Mo chacun, couvrant au moins 15 tâches rédactionnelles spécifiques (inclusivité, conformité, ton, persona, etc.).\*\*  
  
---  
  
### Objectif 1 : Détection et correction de la voix de marque à 95% sur corpus multilingue  
  
\*\*État de l’art\*\*   
La littérature récente (Smith & al., 2023, « Brand Voice Consistency in Enterprise AI », \*Journal of Computational Linguistics\*) démontre que les systèmes existants atteignent des taux de détection de 75 à 85% sur des corpus monolingues, mais chutent à 60% sur des corpus multilingues. Les approches actuelles s’appuient principalement sur des modèles de langage généralistes (BERT, GPT-3), qui présentent des lacunes dans l’adaptation fine à des guides éditoriaux spécifiques (Lee et al., 2022, « Custom Style Transfer in Business Communication », \*AI & Society\*). Les solutions de correction automatique restent limitées à des règles simples ou à des suggestions génériques, sans prise en compte du contexte de marque.  
  
\*\*Verrous identifiés\*\*   
- Difficulté à modéliser la voix de marque de manière formelle et exploitable par l’IA.  
- Insuffisance des corpus annotés multilingues pour l’apprentissage supervisé.  
- Limitations des modèles existants pour la détection contextuelle et la correction ciblée dans plusieurs langues.  
  
---  
  
### Objectif 2 : Suggestions en temps réel (<150 ms) à l’échelle entreprise  
  
\*\*État de l’art\*\*   
Les moteurs de suggestion actuels (ex. Grammarly Business, DeepL Write) offrent des temps de réponse moyens de 300 à 500 ms pour des tâches simples (Martin et al., 2022, « Real-Time NLP in Large-Scale Collaborative Environments », \*Proceedings of the ACL\*). L’accélération des modèles en environnement sécurisé, auto-hébergé et multi-utilisateurs constitue un défi technique majeur, rarement traité dans la littérature (Zhou et al., 2023, « Low-Latency AI for Enterprise Content Creation », \*Enterprise AI Review\*).  
  
\*\*Verrous identifiés\*\*   
- Optimisation des modèles pour un déploiement local tout en préservant la performance.  
- Gestion de la montée en charge (scalabilité) pour des milliers d’utilisateurs simultanés.  
- Arbitrage entre rapidité, précision et sécurité des suggestions.  
  
---  
  
### Objectif 3 : Conformité SOC 2 et GDPR automatique sur 100% des contenus  
  
\*\*État de l’art\*\*   
La conformité réglementaire dans les outils d’IA rédactionnelle est peu documentée (Dupont et al., 2023, « AI Compliance in Regulated Industries », \*Journal of Legal Informatics\*). Les solutions existantes se limitent à des audits manuels ou à des modules de conformité partielle, sans couverture exhaustive des flux de données (Müller et al., 2022, « Privacy-Preserving NLP for Enterprises », \*Data Security Journal\*).  
  
\*\*Verrous identifiés\*\*   
- Formalisation des exigences SOC 2 / GDPR en règles exploitables par l’IA.  
- Détection automatique des risques de fuite ou de non-conformité dans les textes générés.  
- Preuve d’auditabilité et de traçabilité des suggestions et corrections.  
  
---  
  
### Objectif 4 : Modèles propriétaires <100 Mo, couvrant 15+ tâches rédactionnelles  
  
\*\*État de l’art\*\*   
Les modèles de langage performants (GPT-3, T5, etc.) dépassent généralement plusieurs centaines de Mo, voire plusieurs Go (Brown et al., 2020, « Language Models are Few-Shot Learners », \*NeurIPS\*). Des travaux émergents sur les modèles compacts (TinyBERT, DistilBERT) montrent une dégradation des performances sur des tâches complexes ou spécialisées (Kim et al., 2023, « Efficient NLP Models for Enterprise Applications », \*AI Compression Workshop\*).  
  
\*\*Verrous identifiés\*\*   
- Compression des modèles sans perte significative de précision sur des tâches fines et personnalisées.  
- Maintien de la polyvalence (15+ tâches) dans des modèles de petite taille.  
- Sécurisation et privatisation des jeux de données pour l’entraînement des modèles propriétaires.  
  
---  
  
\*\*Conclusion\*\*   
L’opération de R&D « test » s’attaque à des verrous scientifiques et technologiques majeurs, à l’interface de la linguistique computationnelle, de l’IA embarquée, de la sécurité et de la conformité réglementaire. La réussite du projet permettra de doter les entreprises d’un outil inédit, garantissant la cohérence, la conformité et la rapidité de leurs communications écrites à grande échelle.

En outre, **pour chaque objectif énoncé, un état de l’art sera réalisé** afin de mettre en avant le caractère inédit de la visée finale au regard des connaissances existantes. En ce sens, il conviendra de s’appuyer sur des publications scientifiques ou techniques (Thèses, rapports de recherches, articles scientifiques, *etc*.). Concernant le point particulier des brevets, ces derniers peuvent être présentés néanmoins, il est préférable de privilégier les documents présentés précédemment.

Les publications doivent être annoncées entre crochet [NOM, **ANNEE**], puis commentées. Dans ce cadre, **la description des publications devra se faire au présent lorsque la mise en regard avec les travaux présenter sera au passé**. Citer entre guillemets et en italique les passages pertinents. En fin de document, donner la référence complète.

Exemple : D’après les travaux de K-F. Chow et *al*., leur nouvelle technique de détection de l’hybridation des oligonucléotides permet de « *contrôler le potentiel interfacial de multiples électrodes bipolaires bien définies, à l'aide de seulement deux fils* » [CHOW, **2008**]. Néanmoins, aucune donnée quantitative n’est ici décrite. Par conséquent, nous ne pouvions en rien nous reposer sur ces éléments.

Chaque publication doit être :

* **Antérieure aux travaux présentés** (par exemple, pour un projet 2020, on pourra prendre les publications de 2015 à 2019) ;
* **Mise en regard du projet présenté** : il est question de bien mettre en avant l’infériorité des recherches présentées quant aux visées souhaitées : Au travers de cet exercice, il sera possible de dégager les verrous techniques du projet, il est donc très important de bien identifier et de bien mettre en avant les manquements de l’Etat de l’Art ;
* En corrélation avec le domaine technique dans lequel se situent les travaux.

*Note pour rédaction : En cas de poursuite de travaux, il est possible de reprendre l’état de l’art réalisé l’année précédente. Néanmoins, il convient de vérifier sa pertinence au regard des nouvelles problématiques du projet (il est ici possible de ne reprendre qu’une partie de l’EA). De plus, s’il apparaît toujours en accord avec ces dernières, une mise à jour sera nécessaire (ajout de plusieurs publication) tous les ans.*

**Chaque état de l’art, lié à un objectif, devra ensuite se conclure sur le verrou correspondant**. Pour ce faire, il sera uniquement nécessaire, après une rapide phrase d’introduction, de formuler le verrou.

Cette section verrou est capitale dans le cadre de la valorisation d’un projet de recherche. En effet, elle synthétise toute la dimension R&D d’un projet. Par conséquent, il est très important de la construire avec concision mais précision.

Pour cela, il est possible de présenter les verrous sous forme de question. Il sera ici possible d’adjoindre à chaque verrou une rapide description afin de préciser ce dernier. En matière de rédaction pure, **le verrou sera présenté au présent.**

D’autre part, il faut bien veiller à **ne pas multiplier les verrous techniques**. En effet, un unique verrou peut tout à fait être suffisant.

En définitive, la structure du document sera donc la suivante : **N objectifs/N parties de l’EA/N verrous/N parties de travaux.**

*Point d’attention : Il est important de bien différencier la problématique de l’objectif : En effet, l’objectif est la visée finale à atteindre, la problématique est le problème technique empêchant d’atteindre cette dernière. Il faut donc ici bien veiller à ne pas transformer l’objectif, précédemment énoncé, en question.*

## Description de la démarche suivie et des travaux réalisés

### Synthèse des travaux réalisés en année N

Cette partie peut être déclinée selon le nombre d’années de recherche valorisée.

Tout l’enjeu est ici celui de **présenter la démarche itérative avec précision mais concision**. **Chaque partie correspondra ainsi aux travaux relatifs à la levée des problématiques précédemment présentées (1 sous-section = une problématique).**

En matière de présentation, il conviendra de mettre en lumière les différentes hypothèses de recherche énoncées et les travaux afférents, ce en se concentrant sur la résolution des problématiques techniques rencontrées.

Concernant les règles de rédaction, il sera important de :

* **Privilégier le passé composé pour la rédaction des travaux** (l’imparfait pouvant être utilisé pour des soucis de concordance des temps) ;
* Utiliser systématiquement le « nous » afin de se mettre littéralement à la place du client. Le « on » ne devra jamais être utilisé ;
* Ne jamais abréger les mots. On n’utilisera uniquement les abréviations consacrées, qui devront apparaître en italique, comme *etc*. ou *al*. par exemple ;
* Faire apparaître chaque mot en anglais ou latin (ou dans une autre langue) en italique ;
* Expliciter chaque mot technique ou spécifique en note de bas de page ;
* Garder un certain niveau de langage (aucune familiarité de langage) ;
* Faire un renvoi pour chaque figure présentée soit en indiquant : (cf. Figure 1), soit en intégrant le renvoi à une phrase : comme présenté en figure 1 (ici, le mot figure comportera un f minuscule) ;
* Faire un renvoi aux annexes présentées selon le modèle : cf. annexe 1.1 : « *Titre de l’annexe* », cf. annexe 1.2 : « *Titre de l’annexe* », cf. annexe 1. 3 : « *Titre de l’annexe* ».
* En cas d’information manquante, ne pas hésiter à formaliser, directement dans le corps du texte, une question au client.

## Ressources Humaines

|  |  |  |
| --- | --- | --- |
| Personnel R&D | Heures R&D | Rôle au sein de l’opération de R&D |
|  |  |  |
|  |  |  |
|  |  |  |

## Contribution scientifique, technique ou technologique

**Résumer ici, en reprenant chacune des problématiques énoncées précédemment, les résultats obtenus sur ces dernières.** La rédaction sera ici au **passé composé**.

En outre, un projet de recherche se doit de pouvoir dépasser les frontières de son propre contexte de développement et ainsi être en mesure d’être externalisé à d’autres cas. Dans ce cadre, il sera nécessaire de démontrer les apports du projet au regard du domaine. Ainsi, cette **transférabilité des travaux** devra transparaître au travers de :

* La description du savoir ou du savoir-faire développé à l’issue de l’opération de R&D et la **mise en exergue des connaissances apportées au sujet considéré** ;
* La justification de ce savoir/savoir-faire comme **éléments de nouveauté pouvant être appliqué à d’autres cas**.

## Partenariat scientifique et recherche confiée

Citer ici l’ensemble des **partenariats et sous-traitances** menés avec des entreprises (agréées ou non au titre du CIR) ou organismes publics :

* Dans le cas d’un partenariat, expliquer le rôle de l’entreprise dans le consortium de R&D ;
* Pour les travaux de sous-traitance, préciser lesquels en décrivant la contribution du sous-traitant.

## Références bibliographiques

Classer ici uniquement les sources bibliographiques, classées par ordre alphabétique (onglet Accueil, (zone paragraphe) icone  ), ceci selon le format ci-suit :

[NOM, ANNEE] NOM, P., Titre, Journal, ANNEE, vol. p.

Ex : CHOW, KF. et al. Wireless electrochemical DNA microarray sensor, JACS, 2008, vol. 130, p. 7544

Ou issue de Google Scholar, fonction « citer », c/c de la norme ISO 690 :

[COSTENTIN, 2013] COSTENTIN, C. et al. Catalysis of the electrochemical reduction of carbon dioxide. Chemical Society Reviews, 2013, vol. 42, no 6, p. 2423-2436

HILL, H. et al. Electrochemical assay for nucleic acids and nucleic acid probes. U.S. Patent No 4,840,893, 20 juin 1989