

Haute Ecole
« ICHEC – ECAM – ISFSC »



Enseignement supérieur de type long de niveau universitaire

Méthodologie d'application de l'intelligence artificielle au reporting CSRD : développement d'un agent intelligent avec NexusGPT dans le cadre d'un projet chez Intys.

Mémoire présenté par :
Lucie WATELET

Pour l'obtention du diplôme de :
Master en gestion de l'entreprise
Année académique **2024-2025**

Promoteur :
Alain EJZYN

Remerciements

Tout d'abord, j'exprime ma profonde gratitude à mon promoteur, Monsieur Alain Ejzyn, professeur assistant à l'ICHEC, pour son encadrement attentif, sa disponibilité et ses conseils avisés, qui ont été déterminants pour la réalisation de ce mémoire.

Je tiens également à remercier mon maître de stage, Loïc Decaux, pour m'avoir initiée à ce magnifique projet ainsi que pour ses conseils et son soutien tout au long de cette expérience.

Ma reconnaissance envers Hippolyte Langellier-Bellevue est immense pour son expertise, ses encouragements, sa patience et sa bienveillance. Nos multiples réunions, avec leurs retours en arrière et réflexions intenses, ont été déterminantes ; sans lui, je n'aurais pas pu réaliser ce projet. J'adresse aussi mes remerciements à Camille Maertens de Noordhout pour ses conseils et son soutien.

Je remercie vivement toutes les personnes qui ont accepté de participer à mes entretiens et ont généreusement partagé leur temps et leurs connaissances :

- Monsieur Pascal Mukile, Head of Audit, Risk & Quality chez Sibelga.
- Madame Christine Vanoppen, ESG Manager chez Lineas.
- Madame Savitri-Aurélie Deshpande, ESG Expert chez Elia.
- Monsieur Martin Tombal, Sustainability Officer chez Sibelga.
- Monsieur Daksh-Sanjay Daga, Sustainability & CSR Reporting Manager chez Alstom.
- Monsieur Olivier Jamin, Sustainability Manager chez Veolia.
- Madame Rivka Mangold, Responsable RSE chez T&S.

Leurs réponses et leurs témoignages ont été d'une grande richesse pour mes recherches.

Je souhaite aussi adresser un remerciement à Assem Chammah, fondateur et CEO de NexusGPT, pour ses conseils et sa patience lors des réunions liées à la création des agents.

Mes remerciements s'adressent également à mes co-stagiaires, Lorraine, Antoine, Louis, Olivier et Benjamin, pour leurs conseils et leur présence tout au long de ce projet, ayant été les témoins directs de son évolution.

Enfin, je tiens à remercier Patrick Marrette pour sa relecture précieuse, qui a grandement contribué à améliorer la qualité de ce mémoire.

Je souhaite également exprimer toute ma gratitude à mon entourage proche : ma maman, mon frère, ainsi que mes amis Emy, Tania, Youna, Alessandra, Inès, Nathan et bien d'autres, pour leur soutien indéfectible et leurs encouragements tout au long de la rédaction de ce mémoire.

Engagement anti-plagiat

Je soussigné, WATELET Lucie, étudiante en Master 2, déclare par la présente que le travail ci-joint respecte les règles de référencement des sources reprises dans le règlement des études en signé lors de mon inscription à l'ICHEC (respect de la norme APA concernant le référencement dans le texte, la bibliographie, etc.) ; que ce travail est l'aboutissement d'une démarche entièrement personnelle; qu'il ne contient pas de contenus produits par une intelligence artificielle sans y faire explicitement référence.

Par ma signature, je certifie sur l'honneur avoir pris connaissance des documents précités et que le travail présenté est original et exempt de tout emprunt à un tiers non-cité correctement.

Date : 12/08/2025



Je soussigné(e), **Lucie Watelet - 190358** (nom + numéro de matricule), déclare sur l'honneur les éléments suivants concernant l'utilisation des intelligences artificielles (IA) dans mon travail / mémoire :

Type d'assistance		Case à cocher
Aucune assistance	J'ai rédigé l'intégralité de mon travail sans avoir eu recours à un outil d'IA générative.	
Assistance avant la rédaction	J'ai utilisé l'IA comme un outil (ou moteur) de recherche afin d'explorer une thématique et de repérer des sources et contenus pertinents.	<input checked="" type="checkbox"/>
Assistance à l'élaboration d'un texte	J'ai créé un contenu que j'ai ensuite soumis à une IA, qui m'a aidé à formuler et à développer mon texte en me fournissant des suggestions.	<input checked="" type="checkbox"/>
	J'ai généré du contenu à l'aide d'une IA, que j'ai ensuite retravaillé et intégré à mon travail.	<input checked="" type="checkbox"/>
	Certaines parties ou passages de mon travail/mémoire ont été entièrement générés par une IA, sans contribution originale de ma part.	
Assistance pour la révision du texte	J'ai utilisé un outil d'IA générative pour corriger l'orthographe, la grammaire et la syntaxe de mon texte.	<input checked="" type="checkbox"/>
	J'ai utilisé l'IA pour reformuler ou réécrire des parties de mon texte.	<input checked="" type="checkbox"/>
Assistance à la traduction	J'ai utilisé l'IA à des fins de traduction pour un texte que je n'ai pas inclus dans mon travail.	
	J'ai également sollicité l'IA pour traduire un texte que j'ai intégré dans mon mémoire.	
Assistance à la réalisation de visuels	J'ai utilisé une IA afin d'élaborer des visuel, graphiques ou images.	
Autres usages		

Je m'engage à respecter ces déclarations et à fournir toute information supplémentaire requise concernant l'utilisation des IA dans mon travail / mémoire, à savoir :

J'ai mis en annexe les questions posées à l'IA et je suis en mesure de restituer les questions posées et les réponses obtenues de l'IA. Je peux également expliquer quel le type d'assistance j'ai utilisé et dans quel but.

Fait à **Bruxelles** (ville), le **12/08/2025** (date)

Signature : **Lucie Watelet - 190358** [Prénom Nom de l'étudiant(e) et matricule]



Table des matières

INTRODUCTION	1
CHAPITRE 1 : L'HISTOIRE DE LA « CORPORATE SUSTAINABILITY REPORTING DIRECTIVE »	2
1. LA RESPONSABILITÉ SOCIALE DES ENTREPRISES (RSE)	2
1.1. ORIGINES ET ÉVOLUTION DE LA RSE	2
2. DÉFINITION ET PRINCIPES DE LA RSE	3
3. CADRES ET NORMES INTERNATIONALES	4
4. LES NORMES ESG	5
4.1. CONTEXTE ET ORIGINE DES CRITÈRES ESG	5
4.2. DÉFINITION ET STRUCTURE DES CRITÈRES ESG	5
4.3. ESG ET RSE : DES CONCEPTS LIÉS MAIS DISTINCTS	6
4.4. L'IMPORTANCE CROISSANTE DES CRITÈRES ESG POUR LES ENTREPRISES ET LES INVESTISSEURS	6
5. REPORTING EXTRA-FINANCIER	7
5.1. DÉFINITION ET CONTEXTE	7
6. LES GUIDES INTERNATIONAUX	8
6.1. GLOBAL REPORTING INITIATIVE	8
6.2. THE SUSTAINABILITY ACCOUNTING STANDARDS BOARD (SASB)	8
6.3. TASK FORCE ON CLIMATE-RELATED FINANCIAL DISCLOSURES (TCFD)	9
7. L'ÉVOLUTION DES RAPPORTS	9
8. CRITIQUES ET LIMITES DU REPORTING EXTRA-FINANCIER	10
9. LA « NON-FINANCIAL REPORTING DIRECTIVE »	10
10. LE CADRE RÉGLEMENTAIRE : CORPORATE SUSTAINABILITY REPORTING DIRECTIVE	12
10.1. INTRODUCTION DE LA CSRD	12
10.2. CADRE INITIAL DE LA CSRD	12
10.2.1. Comparaison NFRD – CSRD	12
10.3. RÈGLES DE PUBLICATION	13
10.4. LES ESRS – EUROPEAN SUSTAINABILITY REPORTING STANDARDS	14
10.5. LA DOUBLE MATÉRIALITÉ	17
10.6. CHANGEMENT DANS LA DIRECTIVE : L'OMNIBUS	18
11. LES ÉTAPES DU REPORTING	19
12. RÉFLEXION CRITIQUE	27
12.1. LIMITES ET CRITIQUES DE LA CSRD	27
12.2. DE LA CONTRAINTE À L'OPPORTUNITÉ : PERCEPTIONS ÉVOLUTIVES DE LA CSRD SUR LE TERRAIN.	28
13. L'INTELLIGENCE ARTIFICIELLE COMME LEVIER D'OPPORTUNITÉS DANS LE REPORTING CSRD	30
CHAPITRE 2 : L'INTELLIGENCE ARTIFICIELLE	31
1. HISTOIRE, CONCEPT ET DÉFINITION	31
2. LES “AI WINTERS”	31
3. MACHINE LEARNING	31
4. DEEP LEARNING	32
5. TRAITEMENT DU LANGAGE NATUREL (NLP)	33
5.1. DOMAINES D'APPLICATION DU NLP	33
6. L'INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE (IAG)	34
7. LES MODÈLES LLM (LARGE LANGUAGE MODELS)	34
7.1. CONCEPT	35

7.2.	LE CŒUR DU LLM : PRÉDIRE LE MOT SUIVANT	35
7.3.	LA « FENÊTRE DE CONTEXTE »	36
7.4.	LES DIFFÉRENTES ARCHITECTURES DE MODÈLES LLM	36
7.4.1.	Les modèles encodeurs	36
7.4.2.	Les modèles décodeurs	37
7.4.3.	Les modèles encodeur-décodeur	37
7.4.4.	Les modèles multitâches et multimodaux	37
7.5.	RISQUES ET LIMITES	37
8.	GENERATIVE PRE-TRAINED TRANSFORMER (GPT)	38
8.1.	HISTORIQUE ET CONCEPTION DES MODÈLES	38
8.2.	LIMITES ET RISQUES DES MODÈLES GPT	39
9.	PROMPT ENGINEERING	40
9.1.	CONCEPT	40
9.2.	CONSEIL DE RÉDACTION DE PROMPT	40
9.3.	TECHNIQUES DE PROMPTING	41
9.4.	CONSIDÉRATIONS ÉTHIQUES DANS LA CONCEPTION DE PROMPTS	42
—	Limiter les biais	42
—	Respect des règles et des limites	42
—	Vigilance face aux informations erronées	42
—	Transparence et usage responsable	43
10.	LES LIMITES DE L'INTELLIGENCE ARTIFICIELLE	43

CHAPITRE 3 : LES AGENTS D'INTELLIGENCE ARTIFICIELLE	47
------------------------------------------------------------	-----------

1.	DÉFINITIONS ET ORIGINES DES AGENTS D'IA	47
2.	LA TYPOLOGIE DES AGENTS D'IA	48
2.1.	LES AGENTS SYMBOLIQUES	48
2.2.	LES AGENTS RÉACTIFS	48
2.3.	LES AGENTS BASÉS SUR L'APPRENTISSAGE PAR RENFORCEMENT	48
2.4.	LES AGENTS AVEC APPRENTISSAGE PAR TRANSFERT ET MÉTA-APPRENTISSAGE	49
2.5.	LES AGENTS BASÉS SUR DES MODÈLES DE LANGAGE DE GRANDE TAILLE (LLM)	49
3.	ARCHITECTURE D'UN AGENT IA	50
4.	LES LLM : LA BASE DES AGENTS INTELLIGENTS	51
5.	QUELQUES APPLICATIONS CONCRÈTES DES AGENTS D'INTELLIGENCE ARTIFICIELLE	52
6.	LIMITES ET DÉFIS DES AGENTS D'INTELLIGENCE ARTIFICIELLE.	54
6.1.	LIMITES TECHNIQUES DES AGENTS	54
6.2.	RISQUES JURIDIQUES, ÉTHIQUES ET IMPACTS SOCIÉTAUX	55
7.	GESTION DES DONNÉES SENSIBLES DANS LE CADRE DU PROJET	55
8.	COMMENT CONCEVOIR UN AGENT INTELLIGENT : ENSEIGNEMENTS CROISÉS DE LA THÉORIE ET DE LA PRATIQUE	57

CHAPITRE 4 : LE DÉVELOPPEMENT D'UN AGENT SUR NEXUSGPT	59
--------------------------------------------------------------	-----------

1.	PRISE EN MAIN DE LA PLATEFORME NEXUSGPT	59
1.1.	LES CONCEPTS CLÉS DANS LA CRÉATION D'UN AGENT	59
2.	CRÉATION DE L'AGENT DÉDIÉ AU REPORTING CSRD	63
3.	RÉFLEXION SUR LA CONCEPTION DE L'AGENT	64
4.	MÉTHODOLOGIE ET DÉMARCHE DE DÉVELOPPEMENT	65
4.2.	ÉLABORATION DES PROMPTS	69
5.	COMMENT UTILISER LES AGENTS ?	72

6. TEST ET RÉFLEXION	73
7. DIFFICULTÉS RENCONTRÉES LORS DU DÉVELOPPEMENT DE L'AGENT	75
8. CONTRAINTES TECHNIQUES DE L'OUTIL ET LIMITES ACTUELLES	76
8.1. CONTRAINTES LIÉES À L'IA ET AUX LLM (CADRE THÉORIQUE APPLIQUÉ)	76
8.2. CONTRAINTES TECHNIQUES SPÉCIFIQUES AU PROJET	77
9. ÉLARGISSEMENT DU CHAMP D'APPLICATION : VERS LE REPORTING ESG	80
10. RETOUR D'EXPÉRIENCE	80
11. ANALYSE CRITIQUE DU PROJET ET DE SES CHOIX MÉTHODOLOGIQUES	81
 CONCLUSION	 83
 BIBLIOGRAPHIE	 84

Liste des figures

<i>Figure 1 - Le modèle des parties prenantes</i>	3
<i>Figure 2 - Set 1 ESRS</i>	15
<i>Figure 3 - Illustration du concept de double matérialité</i>	17
<i>Figure 4 - Exemple de matrice de double matérialité</i>	22
<i>Figure 5 - Tokenization and Numerical Representation.</i>	35
<i>Figure 6 - Schéma d'une approche progressive de la réalisation d'un prompt.</i>	41
<i>Figure 7 - Cadre conceptuel d'un agent basé sur un modèle LLM</i>	51
<i>Figure 8 - Certifications de sécurité de la plateforme Nexus</i>	57
<i>Figure 9 - Options de sélection de modèles sur l'interface NexusGPT.</i>	60
<i>Figure 10 - Extrait du tableau des exigences de divulgation des ESRS</i>	66
<i>Figure 11 - Table des matières du standard ESRS E1 « Climate Change »</i>	67
<i>Figure 12 - L'interface des agents configurés sur la plateforme NexusGPT</i>	68
<i>Figure 13 - Liste des compétences (GPT-Tasks) assignées à un agent.</i>	69
<i>Figure 14 - Illustration du prompt final utilisé pour la section E1-1 (CSRD)</i>	71
<i>Figure 15 - Génération automatique de la section G1-6 Payment practices par l'agent gouvernance</i>	72
<i>Figure 16 - Évaluation du niveau de conformité de la section G1-1 avec les exigences de l'ESRS.</i>	74

Acronymes

CSRD	Corporate Sustainability Reporting Directive
RSE	Responsabilité sociale des entreprises
ISO	International Organization for Standardization
OIT	Organisation internationale du Travail
OCDE	Organisation de coopération et de développement économiques
GRI	Global Reporting Initiative
SASB	Sustainability Accounting Standards Board
SICS	Sustainable Industry Classification System
TCFD	Task Force on Climate-related Financial Disclosures
FSB	Conseil de stabilité financière
NFRD	Non-Financial Reporting Directive
ESRS	European Sustainability Reporting Standards
EFRAG	European Financial Reporting Advisory Group
IRO	Impacts, risques et opportunités
ISSB	IFRS Sustainability Disclosure Standards
VSME	Voluntary Standard for Micro and Small Enterprises
NLP	Natural Language Processing
IA	Intelligence artificielle
LLM	Modèle de langage de grande taille
RGPD	Règlement Général sur la Protection des Données
AI Act	Artificial Intelligence Act (réglementation européenne sur l'IA)
CLM	Causal Language Modeling
GES	Gaz à effet de serre

Introduction

Aujourd’hui, l’intelligence artificielle transforme profondément notre manière de travailler. Elle permet d’automatiser des tâches longues, répétitives, parfois fastidieuses, en les rendant plus rapides et plus efficaces. Mais peut-elle aussi intervenir dans un domaine plus sensible : celui de la conformité réglementaire ? Peut-on confier à une IA la responsabilité de produire un rapport structuré, normé et juridiquement engageant, en respectant l’ensemble des exigences réglementaires ?

Cette interrogation prend une résonance particulière dans le contexte actuel, où les exigences en matière de transparence et de reporting s’intensifient au sein des entreprises. En Europe, la directive CSRD (Corporate Sustainability Reporting Directive), adoptée en 2022, impose aux entreprises dépassant certains seuils de publier des rapports détaillant leurs impacts sociaux et environnementaux selon les normes ESRS (European Sustainability Reporting Standards). Cette directive remplace la NFRD (Non-Financial Reporting Directive) de 2014, jusque-là plus proche d’un cadre de bonnes pratiques (European Commission, 2024).

À titre d’illustration, le rapport annuel intégré 2024 du groupe Elia comporte 148 pages consacrées à la partie relative au reporting de durabilité conforme à la CSRD (Elia Group, 2024). Ce volume démontre non seulement l’ampleur d’un tel exercice, mais aussi le niveau de structuration attendu de la part des entreprises.

C’est dans ce contexte qu’un besoin client a émergé au sein du cabinet Intys, et plus particulièrement dans sa branche durabilité, Intysify. Intys collabore avec NexusGPT, une entreprise spécialisée dans la création d’agents IA via une plateforme dédiée. L’idée s’est alors imposée : pourquoi ne pas concevoir un agent d’intelligence artificielle capable de générer un rapport CSRD conforme aux normes ? C’est à cette problématique que j’ai été confrontée durant mon stage, en pilotant ce projet de conception d’agent IA de bout en bout.

Ce mémoire est structuré en 4 chapitres. Les 3 premiers posent les bases théoriques nécessaires à la compréhension du projet. Le premier présente le cadre réglementaire de la directive CSRD, le deuxième explore les fondements et enjeux de l’intelligence artificielle, et le troisième décrit les principes de conception d’un agent IA. Le quatrième chapitre est consacré à l’étude de cas, en présentant le projet mené chez Intys : le fonctionnement de la plateforme NexusGPT, la méthodologie de développement adoptée et les résultats obtenus avec la création de 4 agents spécialisés.

Cette organisation vise à accompagner progressivement le lecteur, en partant des concepts théoriques clés pour que, lorsqu’il abordera le cas pratique, il dispose de toutes les bases nécessaires pour comprendre les principes de conception des agents développés.

Chapitre 1 : L'histoire de la « Corporate Sustainability Reporting Directive »

1. La responsabilité sociale des entreprises (RSE)

1.1. Origines et évolution de la RSE

La Responsabilité Sociale des Entreprises (RSE), désignée dans le monde anglo-saxon par l'acronyme « CSR » (Corporate Social Responsibility), trouve ses origines dans le contexte nord-américain. Bien que souvent perçue comme un phénomène contemporain, la RSE remonte aux années 1950, période charnière dans l'évolution du capitalisme américain. Cette époque est marquée par l'essor des grandes entreprises et la transformation progressive de leur statut organisationnel (Acquier & Gond, 2007 ; Jbara, 2017).

Dans ce contexte de mutation, la décennie 1950 est également traversée par une montée des revendications sociales, telles que la liberté d'expression et l'exigence croissante d'une légitimité sociale pour les entreprises. Ainsi, la direction des entreprises, historiquement centrée sur les intérêts exclusifs des actionnaires, enclenche une transition vers une gouvernance élargie prenant en compte d'autres groupes d'intérêt. Ce changement marque une redéfinition du rôle de l'entreprise, qui ne se limite plus à la satisfaction des intérêts actionnariaux. Il s'inscrit désormais dans une logique de responsabilité élargie à l'ensemble des parties prenantes, y compris la collectivité dans son ensemble (Acquier & Gond, 2007).

Le tournant décisif intervient en 1953 avec la publication de l'ouvrage de Howard R. Bowen, *Social Responsibilities of the Businessman*, considéré comme fondateur du champ académique de la RSE (Jbara, 2017).

Cette vision est contestée en 1970 par Milton Friedman dans son article *The Social Responsibility of Business is to Increase its Profits*, publié dans le *New York Times*. Friedman y soutient que seules les personnes physiques peuvent assumer une responsabilité morale, et non les entreprises en tant qu'entités. Il affirme ainsi que les dirigeants n'ont d'autre obligation que de maximiser les profits des actionnaires (Friedman, 1970).

En réaction à cette approche actionnariale, la théorie des parties prenantes, développée par R. Edward Freeman dans *Strategic Management: A Stakeholder Approach*, publié en 1984, propose une vision élargie de la gouvernance. Celle-ci met en avant l'idée que l'entreprise doit intégrer les droits et intérêts de l'ensemble des acteurs avec lesquels elle interagit (Mullenbach-Servayre, 2007).

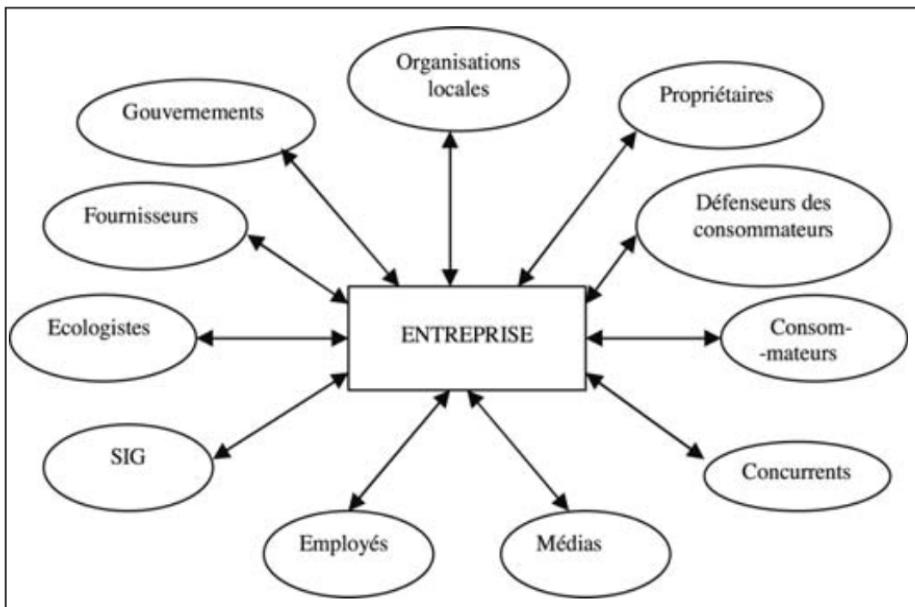


Figure 1 - Le modèle des parties prenantes

Source : D'après Freeman, extrait de Mullenbach-Servayre (2006).

Les décennies 1970 et 1980 sont marquées par une sensibilisation croissante aux impacts sociaux et environnementaux de l'industrie, renforcée par la mondialisation et l'essor des médias de masse (Attarça & Jacquot, 2005). Plusieurs catastrophes majeures alimentent cette prise de conscience. Parmi elles figurent la fuite de dioxine à Seveso, l'explosion de l'usine de Bhopal et les grandes marées noires survenues durant cette période (Connaissance des Énergies, 2021 ; Novethic, 2025 ; Encyclopédie d'histoire numérique de l'Europe, 2025).

Ces événements, associés aux critiques et aux positions théoriques de penseurs tels que Friedman et Freeman, ont mis en évidence les débats sur les dérives d'une industrialisation non encadrée. Ils ont également favorisé l'émergence de mouvements sociaux exerçant une pression croissante sur les entreprises pour qu'elles adoptent des pratiques plus responsables.

2. Définition et principes de la RSE

La responsabilité sociale des entreprises a été définie au fil du temps par divers auteurs. Howard R. Bowen, l'un des pionniers du concept, la décrit comme le devoir des dirigeants d'agir selon les valeurs jugées souhaitables par la société (Bowen, cité dans Acquier & Gond, 2007, p.6). Dans cette continuité, Archie B. Carroll distingue 4 types de responsabilités : économiques (générer du profit), légales (respecter la loi), éthiques (alignement avec les valeurs morales) et philanthropiques (actions volontaires au bénéfice de la société) (Carroll, 1979, p. 500). Il définit ainsi la responsabilité sociale des entreprises comme :

« *The social responsibility of business encompasses the economic, legal, ethical, and discretionary expectations that society has of organizations at a given point in time* » (Carroll, 1979, p. 500).

Donna J. Wood (1991) complète cette approche en soulignant trois principes fondamentaux. Le premier est la légitimité du rôle social de l'entreprise. Le deuxième concerne sa responsabilité envers les effets de ses activités. Enfin, le troisième principe porte sur la discréption managériale permettant d'intégrer ces enjeux à sa gouvernance (Koutaya & Koutaya, 2020 ; Wood, 1991).

L'idée de responsabilité des entreprises vis-à-vis de la société a évolué progressivement pour s'imposer comme une orientation stratégique à l'échelle européenne à partir des années 2000. Dans son *Livre vert* publié en 2001, la Commission européenne affirme que la RSE consiste en « *l'intégration volontaire des préoccupations sociales et écologiques dans les activités commerciales et les relations avec les parties prenantes* » (Commission européenne, 2001, p. 4).

En 2011, elle redéfinit la RSE comme « *la responsabilité des entreprises vis-à-vis des effets qu'elles exercent sur la société* » (Commission européenne, 2011, p. 7). Cette redéfinition élargit son champ aux droits humains, à l'égalité, à la lutte contre la corruption et à la gouvernance durable. Elle insiste également sur l'importance d'un dialogue structuré avec les parties prenantes et fait de la RSE un levier stratégique intégré aux politiques internes de l'entreprise (Commission européenne, 2011, pp. 8-9).

3. Cadres et normes internationales

À l'échelle internationale, plusieurs cadres de référence guident les entreprises dans leur engagement sociétal. Les Principes directeurs de l'OCDE (Organisation de coopération et de développement économiques) pour les entreprises multinationales abordent des enjeux tels que l'environnement, les droits humains et la lutte contre la corruption. Ils encouragent également une contribution active au développement durable (OCDE, 2023).

Un autre cadre de référence majeur est le Pacte mondial des Nations Unies. Il propose dix principes issus de grandes déclarations internationales, relatifs aux droits de l'homme, aux normes du travail, à l'environnement et à la gouvernance éthique (Pacte Mondial Réseau France, n.d.).

De son côté, la Déclaration de principes tripartite de l'Organisation internationale du Travail (OIT), sur les entreprises multinationales recommande des pratiques responsables en matière d'emploi, des modalités de travail et des interactions professionnelles (Organisation internationale du Travail, n.d.).

Enfin, la norme ISO 26000, publiée en 2010, offre un cadre non certifiable pour intégrer la RSE autour de 7 thématiques clés : gouvernance, droits humains, travail, environnement, éthique, consommateurs et développement local (Afnor Groupe, n.d. ; ISO, n.d.).

L'ensemble de ces référentiels fournit une base structurée pour aider les organisations à répondre aux attentes croissantes en matière de durabilité et de responsabilité (Commission européenne, 2011).

4. Les normes ESG

Face aux attentes croissantes des parties prenantes en matière de transparence et de responsabilité, les pratiques de RSE ont progressivement évolué vers une approche plus structurée et mesurable. Ce tournant décisif s'est opéré dans les années 2000, avec l'émergence des critères ESG, marquant une nouvelle étape dans l'évaluation de la performance durable des entreprises (Passas, 2024).

4.1. Contexte et origine des critères ESG

L'acronyme ESG trouve ses racines dans le mouvement de l'investissement socialement responsable (ISR), qui a pris de l'ampleur dans les années 1980-1990. Initialement motivé par des valeurs éthiques liées à des convictions religieuses, cet investissement s'est progressivement orienté vers des préoccupations sociétales, environnementales et de gouvernance d'entreprise. Il s'agissait notamment de se détourner d'investissements jugés immoraux – tels que ceux dans l'industrie de l'armement ou du tabac – pour privilégier des placements alignés sur des valeurs de justice sociale et environnementale (Townsend, 2020).

C'est au début des années 2000 que les critères ESG gagnent en légitimité, notamment avec la publication du rapport « *Who Cares Wins* », initié par le Secrétaire général des Nations Unies dans le cadre du Pacte mondial. Ce rapport souligne que les considérations environnementales, sociales et de gouvernance sont essentielles à la qualité de la gestion d'une entreprise, et qu'elles peuvent contribuer à accroître la valeur actionnariale à long terme (United Nations Global Compact, 2004). Dès lors, les critères ESG s'imposent comme un outil permettant de renseigner les parties prenantes sur la performance des entreprises en matière de développement durable (European Commission, n.d.).

4.2. Définition et structure des critères ESG

Le terme « ESG » regroupe 3 grands piliers :

- Environnement (E) : ce pilier évalue l'impact des entreprises sur l'environnement, incluant la lutte contre le changement climatique, la déforestation, la gestion des déchets, la réduction des émissions de gaz à effet de serre, la pollution de l'eau et la préservation des ressources naturelles (Billio et al., 2021).

- Social (S) : il concerne la manière dont les entreprises gèrent leurs relations humaines et sociales. Cela inclut les conditions de travail, l'égalité salariale, la sécurité des employés, la protection des droits de l'homme, ainsi que les relations avec les fournisseurs (Ayming, 2024 ; Billio et al., 2021).
- Gouvernance (G) : ce pilier concerne la structure de direction de l'entreprise, la transparence des décisions, la composition et le fonctionnement du conseil d'administration, ainsi que les mécanismes de lutte contre la corruption (Ayming, 2024 ; Kain, 2025).

4.3. ESG et RSE : des concepts liés mais distincts

Bien que les termes ESG et RSE soient souvent utilisés de manière interchangeable, notamment dans les entreprises, il est essentiel de distinguer leurs spécificités (Ayming, 2024 ; Kaźmierczak, 2022). La responsabilité sociétale des entreprises constitue un cadre global qui incite les organisations à intégrer des considérations sociales, environnementales et économiques dans leur stratégie. Elle s'adresse principalement aux parties prenantes internes et vise à promouvoir une culture d'entreprise responsable. À l'inverse, les critères ESG sont des indicateurs d'évaluation externe conçus pour aider les investisseurs à évaluer les risques et opportunités associés aux pratiques d'une entreprise (Ayming, 2024). Ces 2 approches partagent un ancrage dans les principes du développement durable (Kaźmierczak, 2022).

Des différences majeures peuvent être relevées (TRC, 2022 cité dans Kaźmierczak, 2022) :

- Mesurabilité : la RSE est difficilement mesurable de l'extérieur, car elle repose sur des initiatives internes et volontaires. Les critères ESG s'appuient sur des données quantifiables, répondant aux exigences de transparence des investisseurs.
- Portée : la RSE est utilisée en interne pour améliorer la culture organisationnelle, tandis que les critères ESG sont exploités par des parties externes (investisseurs, agences de notation, régulateurs) pour l'analyse et la décision.
- Communication : les critères ESG visent à démontrer de manière factuelle les efforts réalisés en matière de durabilité ; la RSE met davantage l'accent sur la communication des engagements de l'entreprise auprès de ses parties prenantes.

4.4. L'importance croissante des critères ESG pour les entreprises et les investisseurs

Les critères ESG connaissent un essor considérable, en réponse à la hausse des attentes sociétales vis-à-vis des entreprises à la recherche de pratiques plus éthiques, ainsi qu'à la demande croissante de produits et services respectueux de l'environnement (Kain, 2025).

Un rapport de KPMG publié en 2019, « *The Numbers That Are Changing the World : Revealing the Growing Appetite for Responsible Investing* », met en lumière l'essor des investissements

durables. Il y est indiqué que 30 000 milliards de dollars d'actifs ont été investis de manière responsable, soit une augmentation de 34 % en seulement 2 ans. Le rapport souligne également une implication accrue des dirigeants : 71 % des PDG interrogés estiment que l'alignement des politiques ESG avec les valeurs des clients constitue désormais une responsabilité personnelle (KPMG Ireland, 2019).

Enfin, le rapport révèle une tendance générationnelle marquée : 49 % des millionnaires de la génération des Millennials fondent leurs décisions d'investissement sur des critères sociaux, et 81 % souhaitent être davantage informés sur les enjeux de l'investissement responsable. Cette évolution témoigne d'une transformation durable des priorités d'investissement, portée par les nouvelles générations (KPMG Ireland, 2019).

Les critères ESG ont ainsi permis d'établir un cadre plus structuré, visant à évaluer de manière claire et cohérente l'engagement des entreprises en matière sociale, environnementale et de gouvernance. Dans cette dynamique, les entreprises ont progressivement adopté des rapports extra-financiers afin de présenter leurs engagements de manière structurée et compréhensible par toutes les parties prenantes.

5. Reporting extra-financier

5.1. Définition et contexte

Le rapport extra-financier consiste à divulguer des informations relatives aux engagements de l'entreprise envers la société, l'environnement et la gouvernance, ainsi qu'aux politiques internes mises en place pour prévenir l'impact de ses activités sur la collectivité. Les premiers rapports extra-financiers ont fait leur apparition dans les années 1970, sous l'impulsion des mouvements sociaux et environnementaux. Ils répondent également à une exigence croissante de transparence et traduisent la volonté des entreprises de démontrer publiquement leur responsabilité sociétale et environnementale (Aureli et al., 2020 ; Fornasari & Traversi, 2024).

Ces rapports permettent aux parties prenantes, internes comme externes, de comprendre et d'évaluer l'impact global des activités de l'entreprise, ainsi que les risques associés. Cette démarche de communication contribue à instaurer une relation durable entre l'entreprise et ses parties prenantes (Fornasari & Traversi, 2024).

Au fil du temps, les scandales économiques et les catastrophes environnementales – cf. section 1.1 consacrée à la RSE – ont accéléré le développement du reporting extra-financier, en mettant l'accent sur les aspects sociaux et environnementaux (Aureli et al., 2020).

Dans sa 13^e édition de l'enquête sur le reporting en matière de durabilité, publiée en 2024, KPMG met en évidence l'importance croissante de ce type de rapport. En Belgique, 88 % des

entreprises déclarent publier des données non financières, soit une hausse de 4 points par rapport à 2022 (Mulkens, 2024).

6. Les guides internationaux

Pendant longtemps, la diffusion d'informations ESG par les entreprises a manqué d'uniformité, chaque organisation adoptant ses propres méthodes et supports de communication. L'absence d'un cadre normatif commun a longtemps conduit les entreprises à aborder les enjeux ESG de manière hétérogène, rendant difficile toute comparaison fiable pour les analystes et parties prenantes (Hales, 2021). Comme mentionné dans la section 1.4, plusieurs référentiels internationaux ont posé les bases d'une responsabilité sociétale renforcée. Leur influence s'est également étendue au domaine du reporting extra-financier, en favorisant l'émergence de normes structurées. Parmi ces initiatives, l'une des plus influentes reste sans conteste la Global Reporting Initiative (GRI) (Fornasari & Traversi, 2024).

6.1. Global Reporting Initiative

Fondée en 1997, à la suite de la marée noire de l'Exxon Valdez, la GRI a pour ambition de développer un langage commun pour le reporting extra-financier. Elle permet aux organisations de communiquer de manière structurée et transparente sur leurs impacts environnementaux, sociaux et économiques. Les normes GRI s'adressent à tous types d'organisations, quels que soient leur secteur, leur taille ou leur localisation. Elles peuvent servir à évaluer les politiques internes ou à définir des objectifs stratégiques. La première version officielle des lignes directrices a été publiée en 2000. Depuis, elles ont fait l'objet de révisions régulières pour s'adapter aux enjeux contemporains (GRI, n.d.).

Selon KPMG (2024b), 77 % des entreprises du G250 (les 250 plus grandes entreprises mondiales) utilisent les normes GRI, un taux stable depuis 2022. Pour le N100 (les 100 plus grandes entreprises par pays), ce taux atteint 71 %, en hausse de 3 points sur 2 ans. Ces chiffres confirment que la GRI demeure le référentiel le plus largement utilisé pour structurer le reporting extra-financier.

6.2. The Sustainability Accounting Standards Board (SASB)

Crée en 2011, la *Sustainability Accounting Standards Board* (SASB) est une organisation sans but lucratif. Elle a pour objectif de guider les entreprises dans la publication d'informations sur les enjeux de durabilité ayant une incidence sur leurs performances économiques, qu'il s'agisse de la liquidité, de l'accès aux financements ou de l'attractivité pour les investisseurs (Hales, 2021 ; SASB, 2023).

Ses normes organisent les informations ESG de manière structurée et harmonisée pour améliorer la lisibilité pour les investisseurs. Reposant sur une logique sectorielle, la SASB a

développé le *Sustainable Industry Classification System®* (SICS®), qui regroupe les entreprises selon des profils de risques et d'opportunités durables communs (Hales, 2021). Ce système comprend 77 industries réparties en 11 secteurs, chacun disposant de son propre ensemble de normes comptables liées à la durabilité (Sustainability Accounting Standards Board, 2018).

6.3. Task Force on Climate-related Financial Disclosures (TCFD)

La TCFD a été créée par le Conseil de stabilité financière (FSB) afin d'aider les entreprises à mieux informer les acteurs financiers sur les enjeux liés au climat (TCFD, 2023). Ses recommandations, publiées en 2017, encadrent la divulgation d'informations financières liées au climat selon 4 thématiques interdépendantes : gouvernance, stratégie, gestion des risques, et indicateurs et objectifs (TCFD, 2017).

La Task Force a été dissoute en 2023, le FSB estimant que la Fondation IFRS pouvait désormais assurer le suivi des progrès en matière de divulgation climatique (TCFD, 2023). Malgré cela, la TCFD a favorisé une adoption mondiale croissante de son cadre.

Selon KPMG (2024b), 72 % des entreprises du G250 appliquent désormais ses recommandations (contre 61 % en 2022), et 43 % du N100 en font de même. Les entreprises privilégient de plus en plus la publication de ces informations dans des rapports ESG ou de durabilité, au détriment des rapports financiers annuels.

7. L'évolution des rapports

Depuis plusieurs décennies, les préoccupations environnementales et sociales ont poussé les entreprises à renforcer la communication sur leurs engagements. Les entreprises adoptent des approches diverses pour présenter leurs informations extra-financières. Certaines intègrent ces éléments directement dans leur rapport annuel, tandis que d'autres préfèrent les regrouper dans des documents séparés, tels qu'un rapport de développement durable ou un rapport spécifiquement axé sur les dimensions environnementales et sociales (Saoussany & Kidaye, 2023).

Ce type de publication volontaire est parfois désigné sous le terme de « reporting sociétal ». Celui-ci regroupe l'ensemble des communications portant sur les engagements et actions environnementaux et sociaux des entreprises, en parallèle ou en complément du reporting financier (Capron & Quairel-Lanoizelée, 2003).

Avec l'émergence des critères ESG, le reporting ESG s'est progressivement imposé comme un cadre de référence central. Il vise à rendre compte des performances et des impacts associés à ces 3 dimensions, tout en identifiant les risques, opportunités et stratégies mises en place en lien avec ces enjeux (Raghavan, 2022).

Dans son rapport de 2022 intitulé « *Big Shifts, Small Steps* », KPMG met en évidence l’importance croissante du reporting ESG à l’échelle mondiale. Selon cette étude, 96 % des entreprises du G250 publient désormais un rapport relatif à leurs pratiques de durabilité ou à leurs performances ESG (KPMG, 2022a).

8. Critiques et limites du reporting extra-financier

S’inspirant de travaux antérieurs, notamment ceux d’Einhorn (2005), de Capron et Quairel (2010) et de Campbell et Slack (2011), Boyer-Allirol (2013) souligne que, même si le reporting extra-financier constitue une bonne initiative, il présente certaines faiblesses importantes. Les exigences concernant le contenu et la quantité d’informations à divulguer manquent notamment de précision et d’uniformité. Toujours selon Boyer-Allirol (2013), bien que le GRI représente une avancée vers la normalisation du reporting extra-financier, son adoption demeure facultative. Les entreprises conservent une grande latitude dans sa mise en œuvre et peuvent se limiter à ne divulguer que les informations jugées favorables à leur image. Ce manque d’uniformité nuit à la comparabilité des données publiées d’une entreprise à l’autre.

Toujours dans son analyse des dispositifs de régulation du reporting extra-financier, Boyer-Allirol (2013) s’inspire des travaux de Mikol (2003), Rivière-Giordano (2007) et Gillet (2010) pour mettre en lumière le manque de vérification et de contrôle de la véracité des informations publiées. En l’absence d’audit externe ou de certification par un expert indépendant, les lecteurs ont peu de raisons de faire pleinement confiance aux données divulguées.

Plus récemment, Fairfax (2023) questionne la montée des rapports ESG et soulève des critiques liées à leur intégrité. Elle s’interroge : comment garantir que l’engagement ESG des entreprises ne relève pas d’un simple discours, voire d’une stratégie de greenwashing ? Certaines entreprises pourraient en effet publier des engagements éloignés de leurs pratiques réelles.

Malgré un renforcement progressif des exigences réglementaires, la publication d’informations extra-financières relève encore largement d’une initiative volontaire. Pour certaines entreprises, elle semble avant tout constituer un levier de positionnement stratégique dans leur environnement, davantage qu’un engagement clair en faveur de la transparence à l’égard des parties prenantes (Boyer-Allirol, 2013).

9. La « Non-Financial Reporting Directive »

Face à l’importance croissante accordée aux enjeux ESG, l’Union européenne a progressivement instauré un cadre réglementaire visant à encadrer la communication des données extra-financières. Elle s’est imposée comme un acteur majeur dans la définition de

standards de transparence, notamment via la directive NFRD (*Non-Financial Reporting Directive*) (Fornasari & Traversi, 2024).

Adoptée en 2014, cette directive constitue un tournant en matière de régulation du reporting extra-financier. Elle impose à certaines grandes entreprises de communiquer des données sur leurs impacts sociaux, environnementaux et de gouvernance, dans le but de rendre ces informations plus comparables et transparentes au sein de l'Union européenne (Aureli et al., 2020).

Selon la directive 2014/95/UE, l'Union européenne précise les catégories d'entreprises concernées : il s'agit des grandes entités d'intérêt public comptant plus de 500 salariés à la clôture de leur exercice. Cela représente environ 11 700 entreprises soumises aux obligations de la NFRD (Commission européenne, 2021). Ces entreprises doivent publier des informations relatives à leurs impacts environnementaux et sociaux, à la gestion de leur personnel. Elles doivent également fournir des informations liées au respect des droits fondamentaux, à la lutte contre les pratiques de corruption, ainsi qu'à la diversité au sein de leurs instances dirigeantes. Pour répondre à ces exigences, elles disposent d'une certaine liberté dans le choix des référentiels qu'elles utilisent pour structurer leur communication. Elles peuvent notamment mobiliser des standards nationaux, des dispositifs de l'Union tels que le système d'audit (EMAS), ou encore s'inspirer de cadres reconnus à l'international tels que ceux proposés par l'ONU, l'OCDE, l'OIT ou la GRI (Union européenne, 2014).

Afin d'accompagner la mise en œuvre de cette réglementation, la Commission européenne a publié en 2017 des lignes directrices non contraignantes, complétées en 2019 par des indications supplémentaires axées sur les informations liées au climat (Commission européenne, 2021).

Cependant, malgré cette avancée en matière de transparence et de responsabilisation des entreprises, la NFRD présente plusieurs limites importantes. Le dispositif législatif ne permettait pas de couvrir adéquatement les attentes des acteurs concernés par les données extra-financières. De nombreuses entreprises fournissaient peu d'informations sur leurs performances en matière de durabilité, malgré l'intérêt croissant des parties prenantes. Lorsqu'elles communiquaient des données, celles-ci manquaient souvent de fiabilité, d'exhaustivité ou de cohérence entre les entités. De plus, un flou subsistait quant aux informations précises attendues de la part des entreprises (Commission européenne, 2021).

Ces limites ont conduit les institutions européennes à repenser le cadre existant, aboutissant à la création de la CSRD (Corporate Sustainability Reporting Directive), adoptée le 21 avril 2021 (KPMG, 2021).

10. Le cadre réglementaire : Corporate Sustainability Reporting Directive

10.1. Introduction de la CSRD

La Corporate Sustainability Reporting Directive (CSRD) est née en réaction aux insuffisances de la NFRD, laquelle ne prévoyait pas de cadre réellement contraignant pour les entreprises. Celles-ci étaient libres de choisir n'importe quel référentiel, national ou international, pour publier leurs informations extra-financières (Pantazi, 2024). La CSRD a été adoptée après 2 années de négociation entre la Commission européenne, le Parlement et le Conseil. Elle vise à accélérer la transition vers une économie durable en renforçant la transparence des entreprises sur leurs actions en matière de durabilité. Elle ambitionne notamment une intégration plus poussée des enjeux sociaux et environnementaux dans la stratégie des entreprises (République française, 2025).

10.2. Cadre initial de la CSRD

10.2.1. Comparaison NFRD – CSRD

La nouvelle directive introduit plusieurs changements majeurs par rapport à la NFRD.

a) Nombre d'entreprises concernées

Tout d'abord, son champ d'application a été considérablement élargi. En plus des entreprises déjà couvertes par la NFRD, la CSRD s'applique aux entreprises qui dépassent au moins 2 des 3 critères suivants (PwC, n.d.-b.) :

- Un bilan total supérieur à 25 millions d'euros,
- Un chiffre d'affaires net supérieur à 50 millions d'euros,
- Un effectif moyen supérieur à 250 salariés.

Outre ces grandes entreprises, la directive s'applique également aux sociétés cotées sur les marchés réglementés européens, à l'exception des micro-entreprises (KPMG, n.d.). Par ailleurs, les filiales peuvent être exemptées si la société mère publie un rapport consolidé contenant toutes les informations requises. Enfin, les entreprises extra-européennes dotées d'une présence commerciale significative dans l'UE (filiale ou succursale) et générant un chiffre d'affaires de plus de 150 millions d'euros sont également visées par la directive. Ce nouveau périmètre a fait passer le nombre d'entreprises concernées d'environ 11 000 (sous la NFRD) à 50 000 sous la CSRD (Pantazi, 2024).

b) **Normes**

Sous la NFRD, les entreprises n'étaient pas tenues de suivre un référentiel spécifique pour la publication de leurs informations. Elles pouvaient s'inspirer de différents cadres internationaux. La CSRD met fin à cette flexibilité en imposant l'utilisation d'un nouveau standard : les European Sustainability Reporting Standards (ESRS). Ces normes ont été créées pour garantir une homogénéité des informations diffusées (Fornasari & Traversi, 2024). Et seront développés dans des sections ultérieures.

c) **Audit**

Autre innovation majeure : l'introduction d'un audit externe obligatoire. Les informations publiées devront être vérifiées pour garantir leur fiabilité (Fornasari & Traversi, 2024).

Pendant une période transitoire de 3 ans, seule une assurance limitée sera requise. Au terme de cette phase, la Commission européenne pourra envisager de relever ce niveau en instaurant une assurance raisonnable, correspondant à des exigences plus approfondies (Baumüller & Grbenic, 2021).

L'assurance raisonnable est l'équivalent d'un audit des états financiers : elle vise à confirmer que les données ont été préparées correctement, qu'elles sont fidèles et reflètent la réalité de l'entreprise. Ce type d'assurance repose sur une démarche approfondie. Il s'agit de bien comprendre le fonctionnement de l'entreprise, d'examiner ses procédures internes, d'identifier les risques éventuels et de vérifier les informations disponibles avant de tirer une conclusion sur la fiabilité des données. En comparaison, l'assurance limitée est moins étendue : elle offre un niveau de confiance modéré, proche d'un examen de cohérence, sans l'intensité d'un audit complet (Shannon, 2024).

d) **Format de publication**

Enfin, la CSRD impose que les rapports soient publiés dans un format électronique standardisé, le XHTML, conformément au règlement européen sur le format électronique unique (ESEF). Les données de durabilité devront être structurées à l'aide du langage XBRL, afin qu'elles puissent être interprétées automatiquement par des logiciels spécialisés (KPMG, 2024a).

10.3. Règles de publication

La mise en œuvre des obligations de publication prévues par la CSRD est progressive. Les premières entreprises concernées seront celles déjà soumises à la NFRD, qui devront publier leur rapport en 2025 sur base des données de l'exercice 2024. Elles seront suivies, en 2026, par les grandes entreprises ne relevant pas de la NFRD mais répondant aux critères de taille. Les

PME cotées devront quant à elles publier leur premier rapport en 2027, sauf si elles choisissent de reporter cette obligation jusqu'en 2028. Enfin, les entreprises de pays tiers réalisant une activité substantielle dans l'UE (notamment un chiffre d'affaires supérieur à 150 millions d'euros) entreront dans le champ d'application en 2029 (EY, 2024).

10.4. Les ESRS – European Sustainability Reporting Standards

Face aux limites de la directive NFRD, notamment l'hétérogénéité des pratiques de reporting, la Commission européenne a engagé une refonte ambitieuse du cadre normatif en matière de durabilité. Cette démarche vise à garantir une meilleure fiabilité des informations extra-financières publiées, en réponse aux attentes croissantes des parties prenantes (Fornasari & Traversi, 2024).

Pour structurer cette nouvelle approche, l'élaboration des normes ESRS a été confiée à l'EFRAG, qui a mobilisé un large éventail d'acteurs – experts, entreprises, chercheurs, investisseurs – afin d'assurer la représentativité et la pertinence du cadre retenu. Les ESRS instaurent un langage commun pour le reporting ESG, fondé sur l'exigence de clarté, de précision et de comparabilité. En rendant les informations sur la durabilité plus claires et comparables, ces normes permettent aux investisseurs de mieux comprendre les actions des entreprises. Elles aident ainsi à diriger leurs financements vers celles qui s'engagent réellement dans la transition écologique et sociale (Fornasari & Traversi, 2024).

Une des spécificités majeures des ESRS réside dans leur ancrage dans le principe de double matérialité. Celui-ci invite les entreprises à articuler leur reporting à la fois autour de leurs impacts sur la société et l'environnement, et des enjeux que ces dimensions peuvent représenter pour leur propre performance (Commission européenne, 2023). Cette notion, centrale dans la logique de la CSRD, fera l'objet d'un développement spécifique dans la section suivante.

Le premier ensemble de normes, appelé ESRS Set 1, regroupe des standards dits « sector-agnostic », c'est-à-dire non spécifiques à un secteur, et applicables à toutes les entreprises concernées par la CSRD. Il a été officiellement adopté par la Commission européenne le 31 juillet 2023 (EFRAG, 2023b ; EY, 2024).

Un deuxième ensemble de normes, appelé ESRS Set 2, devait initialement suivre en juin 2024. Il concerne des secteurs spécifiques tels que l'énergie (pétrole et gaz), le transport routier ou encore l'industrie textile, ainsi que des normes pour les entreprises de pays tiers (EY, 2024). Toutefois, la Commission a annoncé l'abandon de cette initiative lors du paquet « Omnibus », qui sera exposé plus en détail dans une section ultérieure.

Le Set 1 introduit des principes et lignes directrices qui influencent profondément la manière dont les entreprises collectent, analysent et diffusent leurs données de durabilité. Ce set couvre 4 grands domaines d'information (KPMG, 2024a) :

- Gouvernance : responsabilité des organes de direction, gestion des processus et supervision des questions de durabilité.

- Stratégie : description du modèle d'affaires, des objectifs stratégiques, des parties prenantes, de la chaîne de valeur, et présentation de l'analyse de matérialité.
- Gestion des impacts, risques et opportunités (IRO) : informations sur la manière dont l'entreprise identifie et gère les impacts de durabilité, et comment ces éléments sont intégrés dans la stratégie.
- Mesures et objectifs : indicateurs et objectifs de performance en lien avec l'évolution de l'analyse de matérialité.

L'architecture complète du Set 1 est illustrée dans le schéma ci-dessous, proposé par Greenomy :

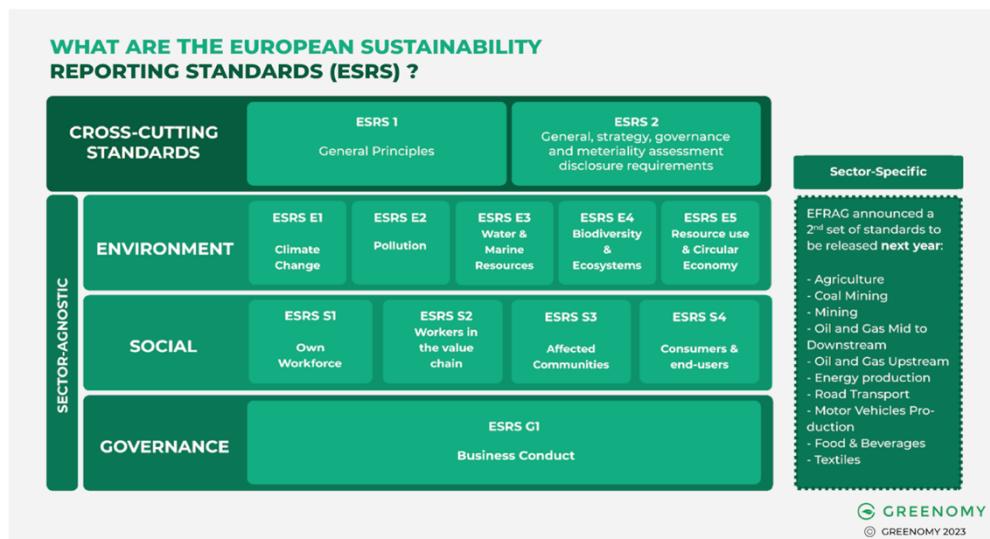


Figure 2 - Set 1 ESRS

Source : Proposé par Gorrin (2023), extrait de « *The European Sustainability Reporting Standards : A Phase-in Journey to Sustainability Alignment* », Greenomy.

Structure des normes ESRS Set 1

Selon l'EFRAG (2023a), la structure du premier ensemble de normes ESRS (*Set 1*) repose sur 2 types de normes : les normes transversales et les normes thématiques.

1. Normes transversales

- *ESRS 1 – General Requirements* : cette norme présente l'architecture générale du cadre ESRS, en introduisant les concepts clés, les principes méthodologiques, ainsi que les règles de base pour organiser et produire les informations de durabilité (EFRAG, 2023a).
- *ESRS 2 – General Disclosures* : elle précise les éléments d'information que l'entreprise doit fournir de manière transversale sur l'ensemble des enjeux jugés pertinents. Elle couvre notamment la gouvernance, la stratégie, la gestion des

impacts, risques et opportunités (IRO), ainsi que les indicateurs et cibles associés (EFRAG, 2023a).

2. Normes thématiques

Ces normes couvrent les différents enjeux du développement durable et sont structurées selon les 3 piliers ESG (EFRAG, 2023a) : environnement, social et gouvernance. Chaque pilier comprend plusieurs sous-catégories (KPMG, 2024a) :

- Environnement :
 - E1 – Changement climatique
 - E2 – Pollution
 - E3 – Ressources hydriques et marines
 - E4 – Biodiversité et écosystèmes
 - E5 – Utilisation des ressources et économie circulaire.
- Social :
 - S1 – Main-d’œuvre propre
 - S2 – Travailleurs de la chaîne de valeur
 - S3 – Communautés affectées
 - S4 – Consommateurs et utilisateurs finaux
- Gouvernance :
 - G1 – Conduite des affaires

Dans une volonté de compatibilité et d’harmonisation internationale, les ESRS ont été conçues pour s’aligner avec d’autres référentiels reconnus, notamment les normes de l’ISSB (IFRS Sustainability Disclosure Standards) et les standards de la GRI (IFRS, 2024). Lors de la publication d’un document conjoint sur l’interopérabilité entre l’EFRAG et la Fondation IFRS, Patrick de Cambourg, président du conseil sur les rapports de durabilité de l’EFRAG, a déclaré :

« On the basis of our joint efforts, we are very happy to be able to issue this very practical guidance on interoperability illustrating in particular how ESRS embed the global baseline approach developed by the ISSB. It demonstrates EFRAG’s commitment to the much needed international convergence of sustainability-related disclosures, on climate and other critical sustainability matters, and our full support to the global momentum in this crucial space » (IFRS, 2024).

En parallèle, l’EFRAG a également renforcé son partenariat avec la GRI, en signant un nouvel accord de coopération après 3 années de collaboration. Cet accord prévoit la publication d’un index d’interopérabilité mettant en lumière l’alignement structurel entre les ESRS et les standards de la GRI (EFRAG, 2023c).

10.5. La double matérialité

Le concept de double matérialité élargit la perspective habituellement centrée sur la performance financière des entreprises. Contrairement à l'approche traditionnelle, dite « matérialité unique », focalisée sur les impacts économiques pour l'entreprise elle-même, la double matérialité englobe également les effets que l'entreprise génère sur la société et l'environnement. Bien que présente dans plusieurs cadres volontaires, cette approche n'a acquis un caractère obligatoire qu'avec l'entrée en vigueur de la CSRD (Belidan & Baghad, 2024). Elle a été formalisée de manière concrète pour la première fois en 2019, à travers le document de la Commission européenne intitulé « *Supplement on Reporting Climate-related Information* » (Adams et al., 2021).

La double matérialité repose sur 2 dimensions complémentaires :

- La matérialité d'impact (*impact materiality*), également appelée approche « *inside-out* », qui s'intéresse aux effets que les activités de l'entreprise peuvent avoir sur la société et l'environnement (Keaveny, 2024 ; PwC, n.d.-a).
- La matérialité financière (*financial materiality*), ou approche « *outside-in* », qui évalue comment les risques et opportunités liés au développement durable peuvent affecter la situation économique et financière de l'entreprise (Keaveny, 2024 ; PwC, n.d.-a).

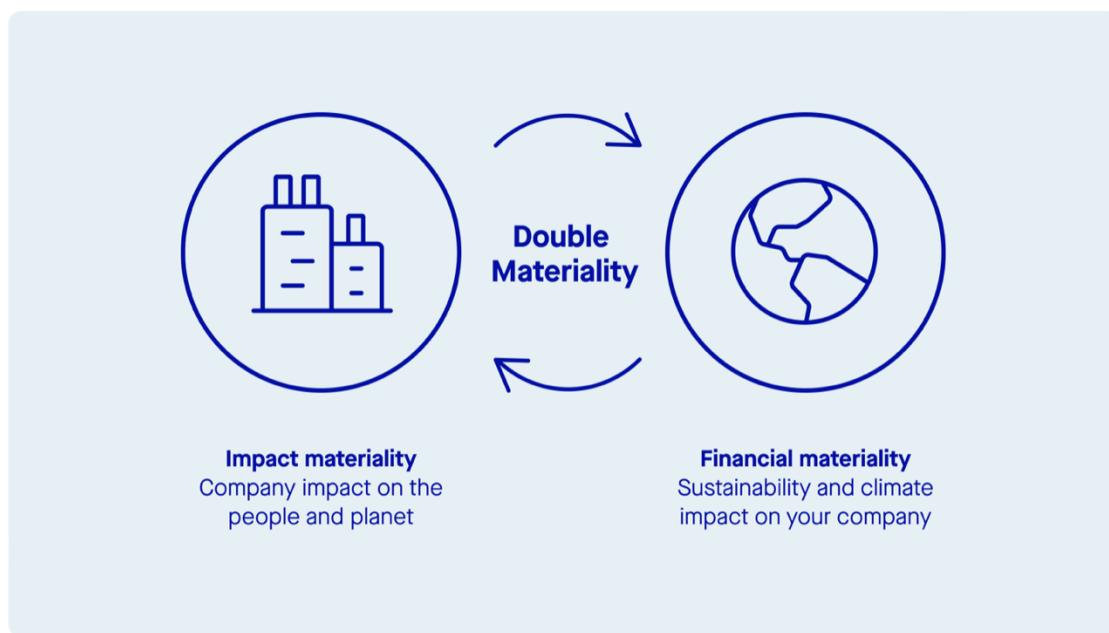


Figure 3 - Illustration du concept de double matérialité

Source : proposé par Sweep (2025), extrait de « *What is double materiality ?* »

Ce principe est aujourd’hui au cœur des obligations de la CSRD et des ESRS, et constitue un changement fondamental dans la manière dont les entreprises sont invitées à évaluer et communiquer leurs enjeux de durabilité.

10.6. Changement dans la directive : l’Omnibus

Le paquet législatif dit « Omnibus », proposé par la Commission européenne le 26 février 2025, vise à alléger les obligations de reporting. L’objectif est de renforcer la compétitivité économique, en particulier pour les entreprises de taille moyenne (Commission européenne, 2025 ; République française, 2025).

L’un des changements majeurs concerne la redéfinition des critères d’éligibilité. Seules les sociétés comptant plus de 1 000 salariés, réalisant un chiffre d’affaires net supérieur à 50 millions d’euros ou disposant d’un total de bilan excédant 25 millions d’euros seraient concernées. Environ 80 % des entreprises initialement visées, y compris certaines des 2 premières phases et la quasi-totalité des PME cotées de la troisième, sortiraient ainsi du périmètre (République française, 2025).

Le calendrier d’application serait également ajusté : les obligations prévues pour 2026 et 2027 seraient reportées à 2028, pour les entreprises concernées par les deuxième et troisième vagues de mise en œuvre (Commission européenne, 2025).

Les entreprises exclues pourraient néanmoins opter pour le standard volontaire VSME, en cours d’élaboration par l’EFRAG. Cette alternative vise les PME souhaitant maintenir une certaine transparence tout en échappant aux obligations strictes de la directive (République française, 2025).

Par ailleurs, les normes sectorielles, prévues pour adapter les ESRS à certains secteurs (énergie, transport, textile), ont été écartées afin d’éviter un alourdissement du reporting (République française, 2025).

Enfin, la révision inclut un ajustement des modalités de certification. L’exigence d’une assurance raisonnable, initialement prévue pour succéder à l’assurance limitée, serait abandonnée afin de contenir les coûts liés à la vérification des rapports (République française, 2025).

Selon Jamin (2025), le projet Omnibus suscite une réaction ambivalente. Il est perçu comme une pause potentiellement bénéfique, permettant à la Commission européenne de réévaluer les attentes formulées à l’égard des entreprises. Toutefois, la réduction drastique du nombre d’entreprises concernées – estimée à 80 % – est interprétée comme un recul, ainsi que comme une atteinte à la crédibilité de l’ambition initiale portée par la directive. Pour une entreprise comme Veolia, la CSRD demeure avant tout une réglementation stratégique, au service de la mise en avant des valeurs de l’entreprise, bien au-delà de la seule contrainte réglementaire.

De son côté chez Sibelga, Mukile (2025) considère l'abandon de l'élaboration des normes sectorielles spécifiques (ESRS sectoriels), tel que proposé par l'Omnibus, comme une décision positive. La CSRD constitue déjà un processus lourd et complexe à mettre en œuvre, et l'ajout de standards encore plus détaillés ne lui semble pas justifié. Il souligne que la double matérialité impose déjà aux entreprises une analyse approfondie de leur chaîne de valeur en lien avec leur secteur d'activité, rendant ainsi superflue la création de normes sectorielles supplémentaires.

11. Les étapes du reporting

En réponse aux besoins d'accompagnement exprimés par les préparateurs de rapports, l'EFRAG a choisi de concentrer ses efforts, dès 2023, sur la publication de documents d'appui à la mise en œuvre des ESRS. Publiés en 2024, ces « *Implementation Guidelines* » couvrent des thèmes techniques tels que la matérialité, la chaîne de valeur, ou encore les jeux de données à reporter (EFRAG, n.d.). Sans valeur normative, ces guides visent à clarifier certains points d'application de la directive. Le guide sur la matérialité rappelle ainsi que : « *No one process would suit all types of economic activities, location(s), business relationships or value chains (upstream and/or downstream) of all the undertakings applying ESRS. An undertaking, based on its specific facts and circumstances, shall design a process that is fit for purpose (...)* » (EFRAG, 2024, p. 26).

Par ailleurs, ces documents abordent des enjeux internes à la CSRD, mais ni l'EFRAG ni la Commission n'ont détaillé le processus de mise en œuvre complète d'un tel rapport. C'est dans cette optique que certains grands cabinets de conseil, tels que KPMG, EY ou PwC, ont développé des cadres méthodologiques permettant d'accompagner les entreprises pas à pas dans l'élaboration de leur rapport. En définitive, le processus de conformité à la directive CSRD et à son contenu représente un véritable défi pour les organisations. Bien que le cadre réglementaire fixe les exigences à respecter, une certaine liberté subsiste quant à la manière dont les entreprises construisent leur démarche de reporting.

La méthodologie du cabinet KPMG est divisée en 5 grandes étapes KPMG (2022b) :

Étape 1 : Impact assessment

La première étape, identifiée par KPMG comme « *impact assessment* », commence par la nécessité, pour les entreprises, de s'approprier pleinement le cadre réglementaire de la CSRD. Le cabinet désigne cette phase sous le terme de « *Regulatory radar* ». Une deuxième composante de cette étape consiste à construire une stratégie de reporting, en définissant la vision de l'entreprise, en analysant sa situation actuelle et en fixant les objectifs visés (KPMG, 2022b).

C'est une étape considérée comme difficile par les entreprises que j'ai pu interroger. Chez Sibelga, Mukile (2025) décrit la CSRD comme trop technique :

« J'ai l'impression que ça a été conçu par des techniciens... et encore. Mais pour moi, la plus grosse difficulté, ça a vraiment été de comprendre cette législation. Parce qu'avant même d'aller voir les gens pour leur demander quoi que ce soit, il fallait qu'on sache nous-mêmes ce qu'on était censés leur demander. Et honnêtement, rien que cette première étape, ça nous a déjà pris beaucoup de temps. »

Chez Elia, c'est la structure du document de l'EFRAG qui a posé problème, jugée peu intuitive et difficile à naviguer en raison de la multiplication des sous-sections, ce qui alourdit la lecture (Deshpande, 2025).

Lineas partage cette difficulté : le référentiel est perçu comme un document massif, truffé d'acronymes, peu accessible pour des personnes non familières des standards réglementaires (Vanoppen, 2025).

Ces témoignages convergent vers une même conclusion : la compréhension initiale de la directive représente une étape chronophage et exigeante, tant sur le plan technique que pédagogique.

Pour y faire face, certaines entreprises comme Alstom ou T&S ont misé sur la formation, afin de s'approprier plus rapidement le cadre réglementaire et de limiter les risques d'interprétation erronée (Daga 2025 ; Mangold, 2025).

Étape 2 : Define Materiality

La deuxième phase de la démarche méthodologique proposée par KPMG (2022b) repose sur 3 composantes complémentaires. Elle débute par une analyse de double matérialité, visant à identifier les thématiques prioritaires à intégrer dans le rapport de durabilité. Cette analyse combine à la fois une matérialité d'impact – c'est-à-dire les effets de l'entreprise sur l'environnement et la société – et une matérialité financière, qui évalue les incidences potentielles des enjeux de durabilité sur la performance de l'organisation. Cette étape s'appuie sur la stratégie définie en amont, les risques identifiés, ainsi que le cadre de gouvernance existant.

Dans un second temps, KPMG (2022b) recommande de définir une stratégie dédiée pour chaque enjeu matériel identifié, en précisant les objectifs visés, les risques et opportunités associés, ainsi que les mesures concrètes à mettre en œuvre. Enfin, des indicateurs de performance (KPI) doivent être déterminés afin d'évaluer l'efficacité des actions entreprises et d'assurer un suivi rigoureux dans le temps.

La mise en œuvre de la double matérialité est une étape particulièrement exigeante, tant en termes de ressources que de technicité. En complément de cette première approche structurée,

PwC (2024) propose également une méthode en 7 étapes permettant de guider plus finement l'analyse de double matérialité.

La première étape, selon PwC (2024), consiste à identifier les parties prenantes et à engager un dialogue avec elles. La CSRD accorde une place centrale aux parties prenantes, les impliquant activement dans l'identification des impacts sociétaux et environnementaux associés aux activités de l'entreprise.

La deuxième étape consiste à recenser les thématiques de durabilité susceptibles d'être significatives pour l'organisation. Ce travail ne se limite pas aux référentiels généraux proposés par les normes ESRS. Il implique aussi une réflexion propre à chaque entreprise, fondée sur son environnement sectoriel, son implantation géographique et la configuration de sa chaîne de valeur (PwC, 2024).

La troisième étape vise à examiner, pour chaque enjeu de durabilité identifié, les effets qu'il peut générer, les menaces qu'il peut faire peser sur l'organisation et les opportunités qu'il peut offrir. Un enjeu ne peut être qualifié de matériel, et donc figurer dans le rapport, que s'il fait l'objet d'une analyse approfondie selon ces 3 dimensions. Cette évaluation nécessite également de préciser l'horizon temporel auquel ces éléments se rattachent – qu'ils soient immédiats, à moyen terme ou plus lointains – ainsi que leur ancrage dans les différentes étapes de la chaîne de valeur. Pour garantir la pertinence et la richesse de l'analyse, il est essentiel de mobiliser à la fois les parties prenantes, des experts thématiques et les différentes unités internes de l'organisation. Cette démarche permet de croiser les regards et d'assurer une compréhension complète des enjeux (PwC, 2024).

Une fois que les enjeux ont été reformulés sous l'angle de leurs impacts, risques et opportunités, il s'agit alors d'en évaluer l'importance avec la plus grande précision possible. Cela implique une quantification fine, conformément aux exigences fixées par les normes ESRS. Cette analyse détaillée permet d'identifier les thématiques réellement significatives, c'est-à-dire celles qui justifient une divulgation dans le cadre du rapport de durabilité (PwC, 2024).

Ensuite, la cinquième étape consiste à examiner les effets financiers liés aux enjeux de durabilité qui ne figurent pas encore dans les états comptables traditionnels. Il s'agit notamment d'évaluer la capacité de l'organisation à conserver l'accès à ses ressources clés et à maintenir des relations pérennes avec ses parties prenantes, dans un contexte de transformation durable (PwC, 2024).

La sixième étape est celle de la cartographie de matérialité. Elle vise à hiérarchiser les IRO identifiés et à retenir les sujets les plus significatifs pour les intégrer dans le rapport (PwC, 2024). À ce stade, l'utilisation d'une matrice de double matérialité s'avère pertinente : elle met en relation, de manière visuelle, les préoccupations sociétales et les incidences économiques, facilitant ainsi l'arbitrage stratégique (République française, 2024).

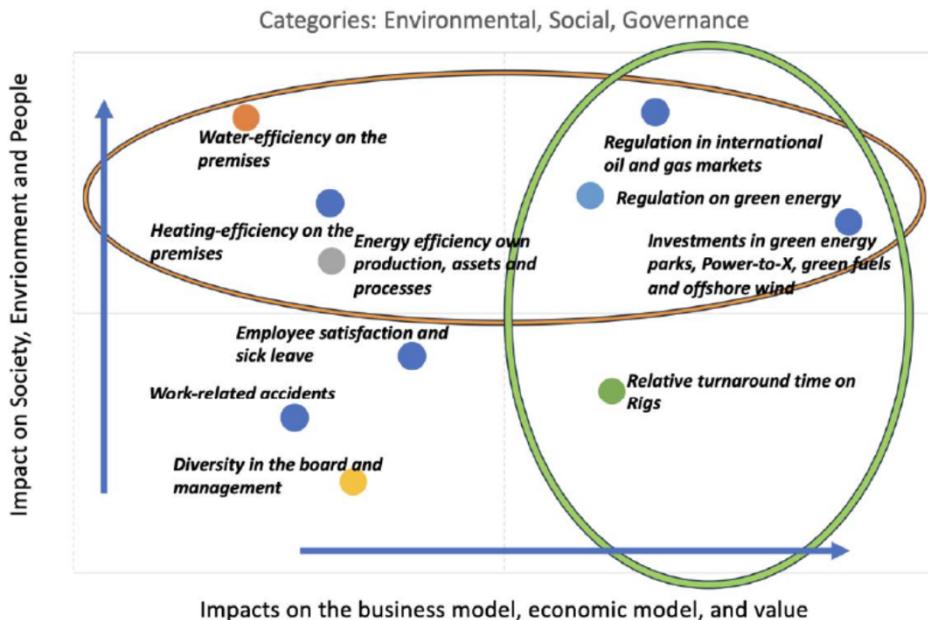


Figure 4 - Exemple de matrice de double matérialité

Source : proposée par Nielsen (2023), extraite de *ESG Reporting and Metrics : From Double Materiality to Key Performance Indicators, Sustainability*, 15(24), 16844.

Enfin, la septième étape consiste à intégrer les thématiques jugées matérielles dans la stratégie globale de l'entreprise. Il s'agit pour l'organisation de décrire les démarches mises en œuvre pour traiter ses principaux impacts, en précisant les leviers d'action mobilisés, les objectifs poursuivis et les indicateurs retenus pour mesurer les progrès réalisés. Cette exigence de transparence implique une planification rigoureuse à long terme, ainsi qu'une coordination interne afin de garantir la cohérence et l'efficacité des politiques déployées (PwC, 2024).

La complexité de cette démarche, tant sur le plan organisationnel que méthodologique, conduit certaines entreprises à solliciter un accompagnement externe. C'est notamment le cas de Sibelga, qui a fait appel à des experts pour conduire cet exercice, considéré comme entièrement nouveau pour l'organisation. Il souligne à ce propos que (Mukile, 2025) :

« *On a fait appel à Deloitte. C'est eux qui ont réalisé l'analyse pour nous, en fait. On a totalement outsourcé ce travail parce que, honnêtement, c'était un sujet complètement nouveau pour nous. On n'avait pas spécialement le temps d'investir là-dedans en interne, donc on leur a confié l'ensemble du processus.* »

De son côté, Lineas a opté pour l'automatisation de son analyse à l'aide d'un outil basé sur l'intelligence artificielle. Après y avoir intégré ses documents internes, l'outil a généré une liste de 118 IRO et produit automatiquement une matrice de double matérialité (Vanoppen, 2025).

Enfin, chez Alstom, la double matérialité constitue un élément structurant de la CSRD. Elle permet, selon lui, de mieux cadrer la stratégie de durabilité et de prévenir les risques de greenwashing en mettant en évidence les véritables enjeux (Daga, 2025).

Étape 3 : Assess Maturity

Après l'analyse de double matérialité, KPMG (2022b) propose une troisième étape intitulée « *Assess maturity* ». Celle-ci a pour objectif d'évaluer le niveau de préparation de l'entreprise en matière de reporting de durabilité, et d'identifier les écarts à combler pour être conforme aux exigences de la CSRD.

La première phase de cette étape correspond à un *gap assessment*, c'est-à-dire une analyse des écarts entre les pratiques de reporting déjà en place au sein de l'entreprise et les obligations spécifiques introduites par la directive (KPMG, 2022b).

Cette étape comprend également une évaluation de l'état de préparation des données utilisées dans le reporting. Pour chaque thématique ou indicateur, elle examine les mécanismes de gouvernance, les procédures internes, les dispositifs de contrôle et les infrastructures de gestion des données (KPMG, 2022b).

Enfin, l'entreprise doit analyser son modèle opérationnel. Cette évaluation porte notamment sur la structuration des processus, la clarté des responsabilités, les capacités internes, l'organisation des systèmes d'information, l'environnement de gouvernance et les règles de gestion existantes (KPMG, 2022b).

Lors de mon entretien avec Sibelga, la question des données manquantes a été explicitement soulevée comme un point de réflexion majeur. L'organisation s'interroge sur l'attitude à adopter en cas d'informations incomplètes :

« *Faut-il assumer publiquement cette absence au risque de nuire à l'image de l'entreprise, ou tout mettre en œuvre pour combler cette lacune, ce qui impliquerait d'engager de nouvelles démarches, d'impliquer davantage les parties prenantes et de réorganiser certains processus internes ?* » (Mukile, 2025).

Étape 4 : Design ESG Reporting Setup

L'avant-dernière étape proposée par KPMG (2022b) concerne la conception du système de reporting. Cette phase vise à structurer l'ensemble des éléments nécessaires à l'élaboration du rapport, tant sur le plan organisationnel que technique.

Une première composante de cette phase consiste à structurer les fondations organisationnelles du reporting. Cela suppose de définir les circuits de validation, de répartir les fonctions-clés entre les équipes concernées. Il convient également de développer les lignes directrices internes en matière ESG, et de mettre en place des dispositifs de contrôle adaptés à la nature non financière des informations traitées (KPMG, 2022b).

À ce stade, les pratiques observées dans les entreprises interrogées varient fortement en fonction des ressources mobilisées.

Chez Sibelga, par exemple, 4 employés sont affectés à temps plein au projet CSRD, dont une personne responsable de l'ensemble du processus. L'entreprise a également instauré un *steering committee* chargé d'expliquer, sur la base de l'analyse de double matérialité, les éléments devant être rapportés (Mukile, 2025).

Lineas, quant à elle, a constitué un comité ESG comprenant 2 représentants des actionnaires. Ce comité joue un rôle de conseil et contribue à structurer la gouvernance. Toutefois, seule une personne est dédiée à la production du rapport CSRD, en collaboration avec les autres départements (Vanoppen, 2025).

Chez Alstom, 3 personnes sont directement responsables du projet CSRD. Elles collaborent avec des interlocuteurs spécifiques selon les thématiques abordées dans les ESRS, avec le soutien d'une équipe plus large mobilisée sur l'ensemble du processus (Daga, 2025).

De son côté, Elia a mis sur pied une équipe entièrement consacrée au reporting, composée de collaborateurs internes et de nouveaux recrutements spécifiques au projet (Deshpande, 2025).

À l'inverse, le groupe T&S disposait déjà d'un service RSE structuré, ce qui lui a permis d'éviter de nouveaux recrutements (Mangold, 2025).

La phase suivante porte sur la structuration du contenu chiffré et narratif du rapport. Elle comprend d'une part la définition des indicateurs de performance, avec les données à collecter et les modalités de calcul, et d'autre part la construction du discours associé pour en expliciter la signification (KPMG, 2022b).

Chez Sibelga, le processus de collecte suit 3 étapes :

« *On a mis en place un programme interne avec différentes étapes bien définies. D'abord, on identifie quelles données doivent être collectées. Ensuite, on vérifie leur disponibilité. Si elles ne sont pas disponibles, on doit alors mettre en place des processus pour les rendre accessibles. On analyse aussi la temporalité des données. Est-ce qu'elles sont disponibles tout au long de l'année ? Est-ce qu'elles ne sortent qu'en début ou en fin d'année ? Sur cette base, on construit notre planning de collecte* » (Mukile, 2025).

Chez T&S, la difficulté réside dans la complexité organisationnelle du groupe :

« (...) *Le challenge chez T&S, c'est que ce n'est pas une seule société, c'est que nous avons beaucoup d'entités. Et parfois, ces entités-là vont elles-mêmes avoir des entités. Donc, c'est ça qui rend la chose un petit peu plus complexe. Donc, en fait, on a chapeauté la collecte des données plutôt au niveau de chacune des entités et pas de chacune des sous-entités* » (Mangold, 2025).

Chez Alstom, la collecte de données s'est révélée particulièrement exigeante, en raison du volume d'informations à compiler et du manque d'outils adaptés. L'entreprise a donc conçu ses propres outils de collecte. La répartition du travail n'a pas suivi une logique par ESRS, mais a été transversale, chaque sujet mobilisant plusieurs départements :

« *I think we were not prepared for the level of data that was required. So in terms of collection and management of data, it became quite complicated for us to make sure that we have a good amount of accounting practices in place (...) each data point, you know, it's like mix and matching different contributors for, let's say, one particular data point. And having 600 plus data points to answer, it was quite complicated that way* » (Daga, 2025).

Elia a opté pour une répartition des ESRS entre les équipes afin de répartir la charge de travail. L'entreprise a identifié la collecte de données comme l'un des aspects les plus complexes du processus. D'une part, en raison de la difficulté à exprimer clairement les données attendues auprès des *business owners*, et d'autre part, à cause de l'incertitude concernant ce que l'auditeur validera réellement. La collecte de données qualitatives, peu habituelle pour l'organisation, a également posé des problèmes d'interprétation (Deshpande, 2025).

Après la collecte des données, les entreprises doivent organiser la rédaction du rapport proprement dit, qui, elle aussi, fait l'objet d'approches variées selon les organisations.

Concernant la rédaction du rapport, les approches varient. Chez Elia, le travail rédactionnel est resté manuel et distribué selon les thématiques ESRS, en suivant la structure de l'EFRAG (Deshpande, 2025). D'autres entreprises comme Lineas et Sibelga ont intégré l'intelligence artificielle via l'outil *Karomia* pour la rédaction (Mukile, 2025 ; Vanoppen, 2025).

Alstom, quant à elle, a opté pour une rédaction humaine, en s'appuyant sur l'outil *Copolit* pour assurer l'uniformité du ton rédactionnel entre les différentes contributions. Contrairement à d'autres, Alstom a aussi fait le choix d'une narration continue et fluide, en s'écartant volontairement de la structure formelle proposée par les ESRS (Daga, 2025).

Pour garantir une approche homogène entre les entités concernées, KPMG (2022b) recommande l'élaboration d'un document de référence interne détaillant la façon dont les KPI ont été construits et mesurés, les règles de consolidation et les méthodes de calcul retenues. KPMG propose aussi de concevoir une architecture technique dédiée au reporting ESG, visant à adapter les outils numériques aux besoins spécifiques du traitement, de la consolidation et de la mise à jour des données.

Étape 5: Group-wide Implementation

La dernière étape décrite par KPMG (2022b) consiste à déployer de manière coordonnée les systèmes, processus et choix méthodologiques élaborés précédemment, afin de garantir leur adoption cohérente au sein de l'ensemble des entités du groupe.

Un aspect clé de cette phase réside dans la préparation à la vérification externe annuelle des informations de durabilité. Cette perspective d'audit systématique soulève différentes réactions de la part des entreprises interrogées, entre préoccupations méthodologiques et contraintes de ressources.

Sibelga perçoit cet audit comme une mesure bénéfique :

« Quand on regarde l'objectif d'un rapport non financier, c'est bien qu'il y ait une vérification. Honnêtement, s'il n'y avait pas d'auditeur, je ne suis pas sûr qu'on aurait mis autant de moyens en place » (Mukile 2025).

D'autres entreprises, comme Elia, adoptent une approche progressive de l'audit, intégrée au fil de l'avancement du processus :

« Mais donc, on a commencé en juillet par l'analyse de la double matérialité et par leur fournir petit à petit notamment des non-financial accounting manuals, donc des documents qui expliquaient comment les données quantitatives étaient rapportées. Ensuite, on a eu des interviews avec eux pour parler des données qualitatives en leur fournissant chaque fois, chapitre par chapitre, donc sur les données quantitatives de l'année écoulée. Évidemment, l'année n'était pas encore écoulée, mais en leur fournissant un peu ce qu'on avait rapporté pour toutes les données qualitatives. Et on leur a fourni tout notre texte fin décembre » (Deshpande, 2025).

Pour Alstom, l'audit est également un processus continu, mené en parallèle de la préparation du rapport. À chaque étape terminée, un entretien a lieu avec l'auditeur, qui formule ses suggestions, donne sa validation ou non. Cette approche est perçue positivement, car elle permet de garantir une cohérence entre les différentes divisions impliquées (Daga, 2025).

Lineas adopte quant à elle une position plus critique à l'égard de cette obligation. Pour l'entreprise, imposer un audit dès la première année, alors qu'aucun standard n'a encore été défini sur ce qui doit être audité, revient à inverser l'ordre logique. Selon elle, la Commission aurait dû commencer par établir les normes. De plus, l'Omnibus a rendu l'assurance raisonnable obsolète, ce qui rend d'autant plus nécessaire la définition de critères clairs permettant aux auditeurs de savoir exactement ce qu'ils doivent faire :

« Mettre un audit obligatoire dès la première année sans qu'il n'y ait un standard établi sur ce qui doit être audité, quelles sont les normes, quelles sont les différentes étapes pour les auditeurs, c'est faire les choses dans le mauvais ordre. La Commission aurait d'abord dû établir les normes avant de mettre un audit obligatoire. De plus, l'Omnibus a rendu l'assurance

raisonnable obsolète, donc établir des normes est une chose possible et permettrait à l'auditeur de savoir exactement ce qu'il doit faire » (Vanoppen, 2025).

12. Réflexion critique

12.1. Limites et critiques de la CSRD

Bien que l'introduction de la CSRD permette de répondre à certaines critiques formulées à l'égard de la directive précédente, plusieurs interrogations et limites subsistent. Des auteurs comme Baumüller et Grbenic (2021) ainsi que Mezzanotte (2023) ont analysé la proposition de la Commission européenne et mis en évidence plusieurs contraintes associées à cette nouvelle directive.

Selon Baumüller et Grbenic (2021), la multiplication des exigences normatives risque de détourner les entreprises de leurs intentions initiales, en les poussant à adopter une approche plus mécanique que stratégique du reporting de durabilité.

Par ailleurs, les entreprises nouvellement soumises à cette réglementation devront adapter rapidement leurs pratiques, ce qui pourrait s'avérer plus difficile que pour celles déjà familières avec la NFRD. Les auteurs soulignent l'ampleur des efforts organisationnels et des investissements requis pour se conformer aux nouvelles exigences, y compris pour les entreprises déjà soumises au dispositif précédent (Baumüller & Grbenic, 2021).

Mezzanotte (2023) ajoute que les moyens dont disposent les entreprises pour la collecte de données peuvent être limités, notamment en raison d'un manque de définition claire des parties prenantes ou d'une chaîne de valeur trop étendue. Certaines entreprises peuvent avoir une chaîne de valeur comprenant plusieurs milliers de fournisseurs, ce qui rend la collecte d'informations particulièrement complexe. Ce manque de moyens peut entraîner des données moins fiables, inexactes ou incomplètes, exposant ainsi les entreprises à un risque d'accusation de greenwashing.

Les entretiens que j'ai menés avec Alstom (Daga, 2025), Elia (Deshpande, 2025), Lineas (Vanoppen, 2025), Sibelga (Mukile, 2025 ; Tombal, 2025), T&S (Mangold, 2025) et Veolia (Jamin, 2025), dans le cadre de mon stage, confirment sur le terrain les limites pratiques identifiées dans la littérature. Ils mettent également en lumière d'autres contraintes observées. Les constats présentés ci-dessous proviennent de l'ensemble de ces entretiens, sauf mention contraire.

Dans un premier temps, la complexité des textes publiés par l'EFRAG a posé un réel problème aux entreprises interrogées. Ces documents sont jugés trop techniques et difficilement accessibles.

De plus, les témoignages confirment les observations de Baumüller et Grbenic (2021) concernant les investissements supplémentaires nécessaires. Certaines entreprises ont dû recruter ou former du personnel, voire créer une équipe spécifique, tandis que d'autres, comme Lineas, évoquent le manque d'effectifs comme un frein majeur. Comme l'explique Vanoppen (2025) : « *It's I, me and myself. (...) So yeah, I'm doing this alone* », illustrant ainsi l'absence de ressources humaines supplémentaires dédiées à cette mission.

Par ailleurs, plusieurs interlocuteurs ont indiqué que la double matérialité représente une nouvelle obligation souvent difficile à comprendre pour les entreprises, et qu'elle demande un temps conséquent pour être correctement traitée. Sibelga, jugeant l'exercice trop complexe, a fait appel à une aide extérieure pour réaliser sa double matérialité (Mukile, 2025). Lineas, de son côté, a eu recours à l'intelligence artificielle pour générer automatiquement son analyse (Vanoppen, 2025).

Concernant la collecte de données, les entreprises soulignent qu'elle constitue une difficulté majeure, ce qui confirme également les analyses de Mezzanotte (2023). Elles rencontrent des obstacles liés à l'indisponibilité de certaines données, ainsi qu'à la nécessité d'en garantir la qualité. Ce manque de fiabilité peut engendrer un risque de greenwashing involontaire, comme l'a également souligné Mezzanotte (2023).

La complexité de certaines structures, comme celle de T&S, rend la collecte encore plus difficile, dans la mesure où les données doivent être extraites de plusieurs entités (Mangold, 2025). Chez Alstom, la taille importante de la chaîne de valeur constitue également un facteur de complexification, en raison du nombre élevé de partenaires impliqués (Daga, 2025).

Une autre contrainte fréquemment évoquée par les entreprises concerne la question de l'audit, qui reste encore relativement floue pour la majorité d'entre elles. Il leur est difficile de déterminer ce qui est précisément attendu par les auditeurs.

Ce constat démontre que, bien que la CSRD constitue une avancée majeure en matière de reporting et représente une véritable opportunité pour renforcer la transparence et harmoniser les standards, son application concrète soulève encore de nombreuses difficultés et limites pour les entreprises.

12.2. De la contrainte à l'opportunité : perceptions évolutives de la CSRD sur le terrain.

Lors de mes entretiens, il m'a paru pertinent de demander aux responsables rencontrés si la mise en place de la CSRD était perçue comme une contrainte ou comme une opportunité.

Du point de vue d'Elia et de Lineas, la CSRD a été perçue dès le départ comme une opportunité.

De son côté, Deshpande (2025) explique que chez Elia, pour l'équipe de durabilité, la directive permet de renforcer le poids du reporting en exigeant que les données extra-financières soient considérées avec la même rigueur que les données financières.

Chez Lineas, Vanoppen (2025) souligne que la CSRD a permis de structurer plus efficacement la stratégie ESG et d'améliorer la transparence vis-à-vis des parties prenantes. Elle insiste également sur l'importance du cadre méthodologique, qui permet à l'entreprise de se comparer à ses pairs et de mieux intégrer les enjeux ESG dans la gouvernance :

« It's putting definitely ESG on the table of the board of directors, the shareholders, in a very explicit way. And that, I think, is a good thing » (Vanoppen, 2025).

Elle précise toutefois qu'elle comprend que certaines entreprises, notamment les plus petites, puissent percevoir la directive comme une contrainte (Vanoppen, 2025).

Cette dernière perception a d'ailleurs été dominante, au moins dans un premier temps, chez les autres entreprises interrogées. Sibelga a d'abord considéré la CSRD comme une charge de travail supplémentaire. Mais en se plongeant dans les exigences de la directive, la perception a évolué. Mukile (2025) explique :

« Au départ, la CSRD, on l'a vraiment perçue comme une obligation en plus (...) Et puis, en tout cas pour moi, après m'être vraiment plongé dans ce qu'était la CSRD, j'ai compris que c'était en réalité une opportunité plus qu'une contrainte. Quand tu commences à saisir l'objectif de la double matérialité, le fait d'être conscient des facteurs externes qui impactent notre business, notre chaîne de valeur... (...) Ça te permet d'identifier les risques, de mieux les maîtriser. Et au final, ça devient un vrai outil. (...) Je suis passé de "c'est une contrainte" à "c'est une opportunité". »

Veolia adopte également une position intermédiaire :

« Le premier réflexe, c'est plutôt un sentiment de contrainte réglementaire. Mais assez rapidement, quand on se plonge dans la matière, on se rend compte que c'est plutôt une opportunité, à un niveau plus global » (Jamin, 2025).

Chez Alstom, le ressenti est également partagé. Daga (2025) parle d'un équilibre entre opportunité et contrainte :

« It's a 50-50 feeling. It's both an opportunity and a constraint. »

Il mentionne notamment la complexité méthodologique de la directive, la quantité de données à collecter (ex. : scopes 1, 2, 3), ainsi que la difficulté de compréhension des textes. En parallèle, il reconnaît que la directive offre un cadre structurant et pousse les entreprises à développer leurs propres solutions (Daga, 2025).

Enfin, T&S partage une lecture similaire. Pour Mangold (2025), la CSRD est perçue comme une contrainte au départ, comme toute nouvelle réglementation, mais devient une opportunité dans la mesure où elle pousse l'organisation dans la bonne direction :

« Alors, c'est toujours une contrainte, au début. Quand il y a une nouvelle chose qui arrive, c'est toujours un peu vécu comme une contrainte. Mais après, c'est aussi une opportunité, parce qu'elle nous pousse dans le bon sens. (...) Il y a certaines choses qu'on ne faisait pas. »

Ces différents témoignages montrent que, bien que la CSRD soit perçue dans un premier temps comme une contrainte, elle tend à être reconstruite comme une opportunité stratégique dès lors que ses lignes directrices sont mieux comprises et intégrées.

13. L'intelligence artificielle comme levier d'opportunités dans le reporting CSRD

L'intelligence artificielle s'impose aujourd'hui comme un levier central de transformation pour les entreprises, notamment en matière d'optimisation des processus et de gain de productivité. Dans le cadre de ce mémoire, les différents entretiens que j'ai menés ont permis d'interroger les professionnels sur l'opportunité d'intégrer l'IA dans la production de rapports de durabilité conformes à la CSRD. Tous s'accordent à reconnaître la grande utilité de cette technologie, en particulier pour automatiser le tri des données, structurer l'information, ou encore accélérer la rédaction des narratifs. Certaines entreprises ont d'ores et déjà recours à des outils d'IA pour les accompagner dans l'analyse de double matérialité, la collecte et la restitution des données, ou encore la génération des contenus requis par les ESRS.

Vanoppen (2025) affirme : « *You can reduce your workload. Here with 90 %* », illustrant ainsi l'ampleur du potentiel de cette technologie dans le contexte du reporting durable.

Ces témoignages confirment que l'intelligence artificielle trouve sa place dans une logique de conformité structurée autour de standards stricts comme ceux de la CSRD. C'est dans cette perspective que s'inscrit également mon projet de stage chez Intys, qui vise à développer un agent capable d'assister les entreprises dans la rédaction de leur rapport de durabilité. Ce projet de stage fera l'objet d'une analyse approfondie dans le quatrième chapitre, précédée d'un cadre théorique permettant de définir les concepts clés liés à l'intelligence artificielle dans le chapitre 2 et aux agents IA dans le chapitre 3.

Chapitre 2 : l'intelligence artificielle

1. Histoire, concept et définition

Bien avant l'apparition officielle du terme « intelligence artificielle », introduit par John McCarthy lors de la conférence fondatrice de Dartmouth en 1956 (Xu et al., 2021), les premières réflexions sur le sujet remontent aux travaux d'Alan Turing dans les années 1950. Considéré comme l'un des fondateurs de l'informatique moderne, il conçoit des dispositifs capables d'exécuter des suites d'instructions suffisamment sophistiquées pour reproduire certaines formes de raisonnement humain (Benani, 2024 ; Hirani et al., 2024). L'événement de Dartmouth marque le début officiel de l'IA en tant que champ scientifique structuré. Il ouvre la voie à de nombreux travaux axés sur l'automatisation du raisonnement, dont les retombées se feront sentir dans des disciplines aussi diverses que les mathématiques, l'informatique ou encore la logique (Xu et al., 2021).

2. Les “AI Winters”

Le parcours de l'intelligence artificielle a été marqué par 2 périodes de ralentissement majeures, désignées comme des « Hivers de l'IA ». Le premier survient à la fin des années 1960, lorsque le fossé entre les espoirs médiatiques, les ambitions parfois irréalistes des chercheurs et les résultats encore limités entraîne un désengagement progressif du soutien institutionnel et financier. Les projets sont freinés et les perspectives du domaine s'assombrissent jusqu'au début des années 1980 (Toosi et al., 2021).

Un renouveau émerge alors avec les systèmes experts : ces programmes, fondés sur des règles dérivées de connaissances expertes, suscitent un fort engouement industriel grâce à leur efficacité dans des domaines ciblés (Toosi et al., 2021). Mais leurs limites – coûts d'entretien élevés, rigidité, manque de flexibilité – conduisent à un nouveau désenchantement dès la fin des années 1980. Le terme même d'« intelligence artificielle » est alors évité par certains chercheurs. Paradoxalement, cette période marque le début d'un tournant, avec l'essor de la statistique appliquée, des capacités de calcul accrues, de l'émergence du web et des données massives (Toosi et al., 2021).

3. Machine learning

Le concept de *machine learning*, ou apprentissage automatique, s'est développé dans les années 1980 et 1990. Il s'agit d'une sous-catégorie de l'intelligence artificielle fondée sur l'auto-apprentissage à partir d'algorithmes utilisant des données préexistantes. Ces algorithmes permettent aux machines de progresser et d'acquérir des connaissances sans intervention humaine (Benani, 2024 ; Sharma et al., 2021).

3 grandes catégories de *machine learning* sont distinguées (Mahesh, 2020 ; Sharma et al., 2021) :

- a) Apprentissage supervisé : méthode reposant sur des données annotées, dans laquelle l'algorithme apprend à associer des entrées à des sorties attendues à partir d'un jeu d'entraînement, puis généralise ses prédictions sur un jeu de test.
- b) Apprentissage non supervisé : méthode dans laquelle l'algorithme identifie des régularités, des structures ou des regroupements dans les données, sans supervision extérieure.
- c) Apprentissage par renforcement : méthode fondée sur des interactions successives avec un environnement, où le système apprend à sélectionner les actions les plus pertinentes en fonction des conséquences observées, dans une dynamique de récompense cumulative.

Aujourd'hui, le *machine learning* transforme en profondeur l'analyse des données massives. Il permet de classifier, prédire, et accompagner la prise de décision dans des domaines variés : santé, ingénierie, finance, psychologie, ou encore sciences sociales (Xu et al., 2021).

4. Deep learning

L'émergence de l'apprentissage profond, *deep learning*, constitue une étape décisive dans l'évolution de l'intelligence artificielle contemporaine. Issu du champ du *machine learning*, ce domaine se caractérise par l'utilisation de réseaux de neurones artificiels, conçus non pas pour reproduire le fonctionnement biologique du cerveau, mais pour optimiser, de manière mathématique, le traitement de données complexes au moyen de structures multicouches (Benani, 2024).

Le qualificatif *profond* renvoie au nombre élevé de couches intermédiaires entre les données en entrée et les prédictions en sortie. Le *deep learning* s'appuie à la fois sur l'apprentissage par l'expérience et sur l'exploitation de jeux de données volumineux, permettant une amélioration progressive des performances au fil des cycles d'entraînement du modèle. Cette approche permet de traiter des volumes d'informations plus complexes, ce qui en fait une méthode plus puissante (Sharifani & Amini, 2023 ; Sharma et al., 2021).

Aujourd'hui, ces technologies sont massivement utilisées dans des domaines variés comme la reconnaissance d'images et de voix, les systèmes de recommandation ou la traduction automatique. Elles sont au cœur de nombreux services du quotidien : assistants vocaux, plateformes de streaming, moteurs de recherche (Sharifani & Amini, 2023 ; Sharma et al., 2021).

Porté par la montée en puissance des capacités de calcul et l'explosion des données, le *deep learning* occupe désormais une place centrale dans des domaines stratégiques tels que la médecine, la conduite autonome ou le traitement du langage naturel (Sharifani & Amini, 2023).

5. Traitement du langage naturel (NLP)

L'ambition de doter les machines de la capacité à comprendre et produire du langage humain constitue l'un des fondements historiques de l'intelligence artificielle (Scarcello, 2019). Le traitement du langage naturel (*Natural Language Processing*, ou NLP), sous-domaine de l'IA, a connu une évolution remarquable grâce aux progrès du *machine learning* et du *deep learning* (Chai et al., 2024). Il regroupe aujourd'hui des techniques telles que la recherche sémantique, la traduction automatique, la reconnaissance d'entités nommées ou l'analyse des sentiments (Abro et al., 2022).

Les premières expérimentations datent des débuts de l'IA, lorsque les chercheurs tentaient de comprendre comment les machines pouvaient manipuler le langage humain. Ces approches ont ensuite été remplacées par des méthodes basées sur les données textuelles, favorisées par l'essor des modèles statistiques dès la fin des années 1980. L'utilisation des probabilités a permis des avancées importantes dans l'analyse automatique du texte, accélérant le développement du NLP (Alabi & Gbemigb, 2025).

Les années 2000 marquent 2 progrès majeurs. La première est l'apparition des *word embeddings*, qui traduisent les mots en vecteurs numériques encodant leurs relations sémantiques. La seconde est l'intégration des réseaux neuronaux dans les systèmes de reconnaissance de la parole et de traduction, offrant une meilleure prise en compte du contexte linguistique (Alabi & Gbemigb, 2025).

L'introduction de l'architecture Transformer en 2017, développée par Vaswani et al. (cité par Alabi & Gbemigb, 2025) a marqué un tournant. En rompant avec le traitement séquentiel traditionnel, ce modèle a permis l'émergence des modèles de langage de grande taille (LLM) et des approches fondées sur des modèles pré-entraînés, désormais omniprésents dans les applications linguistiques (Alabi & Gbemigb, 2025).

5.1. Domaines d'application du NLP

Le traitement automatique du langage naturel (NLP) a profondément transformé et optimisé divers secteurs d'activité grâce à la diversité de ses techniques.

Les assistants vocaux et moteurs de recherche représentent une application courante et bien connue du NLP, tels que Siri, Google Search ou encore Alexa. Ils reposent sur des modèles avancés de NLP dans le but d'analyser les besoins exprimés par les utilisateurs et d'y apporter une réponse pertinente selon la situation (Alabi & Gbemigb, 2025).

Dans le secteur médical, les techniques de NLP sont utilisées pour explorer les données contenues dans les dossiers cliniques, afin d'aider les professionnels à poser un diagnostic fondé sur des éléments textuels fiables. Dans le domaine du service client, l'automatisation des interactions par le biais de chatbots a permis de mieux répartir la charge de travail, en traitant une partie des requêtes sans intervention humaine (Alabi & Gbemigb, 2025).

Dans le secteur éducatif, les modèles de NLP ont favorisé l'amélioration de l'apprentissage des langues. En finance, les techniques d'analyse de sentiment permettent d'évaluer l'impact des mots employés dans les rapports de gestion. Dans le marketing, le NLP est mobilisé pour décrypter les réactions des consommateurs face aux contenus en ligne et affiner la compréhension de leur perception des marques (Abro et al., 2022).

6. L'intelligence artificielle générative (IAG)

Contrairement à l'intelligence artificielle prédictive, fondée sur l'analyse de données passées pour anticiper des événements futurs, l'intelligence artificielle générative (IAG) repose sur la production de contenus inédits à partir de vastes corpus de données. Elle utilise notamment une *fenêtre de contexte*, qui lui permet de maintenir temporairement une mémoire des éléments précédents, assurant ainsi une certaine continuité dans les contenus générés (Gupta et al., 2024).

L'IAG est capable de produire une diversité de contenus multimodaux – qu'il s'agisse de textes, d'images, de sons ou de vidéos – et son utilisation s'étend à un nombre croissant de secteurs. En 2021, le marché mondial de cette technologie était évalué à 7,9 milliards de dollars, avec des projections atteignant 110 milliards de dollars à l'horizon 2030. Les domaines de la santé et de l'éducation en sont particulièrement friands : on y recourt par exemple pour améliorer la radiologie pédiatrique, affiner les processus de diagnostic ou concevoir des outils d'apprentissage innovants. Ces avancées s'appuient principalement sur les modèles de langage de grande ampleur (LLM), qui constituent aujourd'hui l'architecture centrale de l'IAG (Gupta et al., 2024).

7. Les modèles LLM (Large Language Models)

Cette section est développée de manière plus approfondie que les précédentes, de façon volontaire. En effet, le projet réalisé dans le cadre de mon stage chez Intys repose sur l'utilisation de la plateforme NexusGPT, un outil spécialisé dans la création d'agents d'intelligence artificielle pour une grande diversité d'objectifs. Ces agents s'appuient sur des modèles de langage de grande taille (LLM), et plus particulièrement sur des modèles de type GPT. Puisque la plateforme mobilise directement cette technologie, il est essentiel d'en comprendre les principes fondamentaux pour saisir le fonctionnement des agents développés. C'est pourquoi cette section propose une analyse détaillée des bases techniques sur lesquelles reposent les LLM.

Le cadre concret de collaboration avec Intys, ainsi que les objectifs et les résultats du projet, seront présentés dans le quatrième chapitre de ce mémoire, consacré à l'étude de cas.

7.1. Concept

Selon Alabi et Gbemigb (2025), l'architecture des transformateurs a été introduite par Vaswani et ses collaborateurs en 2017 dans leur article « *Attention Is All You Need* », marquant une rupture dans le traitement automatique du langage. Ce modèle repose sur un mécanisme d'auto-attention, qui permet de traiter l'ensemble des éléments d'une séquence en évaluant leurs relations contextuelles, sans passer par une structure strictement séquentielle. D'après ces auteurs, cette innovation a considérablement amélioré les performances des systèmes dans diverses tâches linguistiques (Alabi & Gbemigb, 2025 ; Chai et al., 2024).

Le mécanisme d'attention permet notamment au modèle de hiérarchiser les informations en fonction de leur pertinence contextuelle, en pondérant les contributions respectives des mots au sein d'une phrase (Ramos et al., 2024).

Les modèles LLM sont généralement pré-entraînés de manière auto-supervisée à partir de vastes ensembles de textes, dans le but de développer une compréhension générale de la structure linguistique. Ce processus varie selon les objectifs assignés et l'architecture adoptée, qu'elle soit de type encodeur, décodeur ou hybride (Moradi et al., 2025). Ces modèles sont conçus pour exécuter une grande variété de tâches, allant de la génération de texte à la traduction, en passant par la synthèse ou la réponse à des requêtes. Leur capacité repose sur un apprentissage statistique permettant d'anticiper les unités linguistiques les plus probables à produire en fonction du contexte immédiat (Hipólito, 2024).

7.2. Le cœur du LLM : prédire le mot suivant

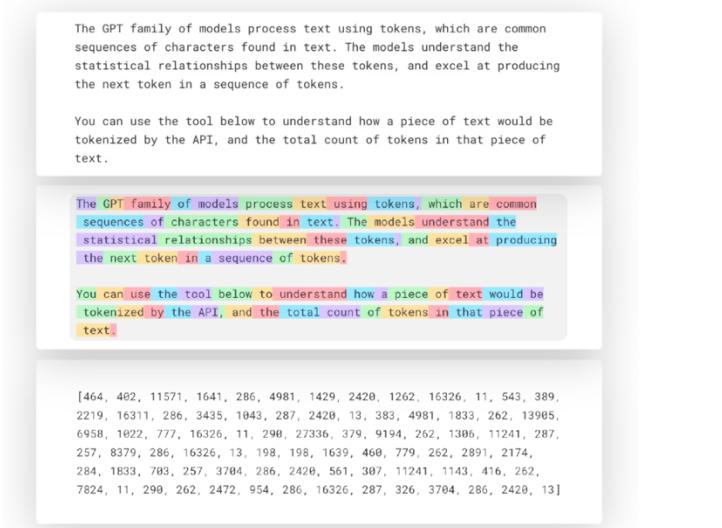


Figure 5 - Tokenization and Numerical Representation.

Source : NexusGPT (2025d), extrait de *Foundational Knowledge*.

Le principe de base des modèles de langage de grande taille consiste à estimer, à chaque étape de génération, quelle unité textuelle a le plus de chances d'apparaître ensuite. Cette unité, appelée *token*, peut correspondre à un mot entier, une partie de mot ou un signe de ponctuation. Avant que le modèle puisse traiter une séquence, celle-ci est découpée selon un procédé appelé *tokenisation*. Chaque token est ensuite représenté sous forme de vecteur numérique, appelé *embedding*, qui contient des indices sémantiques et grammaticaux essentiels à l'analyse contextuelle par le modèle (NexusGPT, 2025d).

7.3. La « fenêtre de contexte »

Les modèles LLM sont limités par un paramètre appelé fenêtre de contexte, aussi désigné comme longueur ou séquence contextuelle. Ce paramètre correspond au nombre maximum de *tokens* que le modèle peut prendre en compte simultanément lors de la génération d'un texte (Hipólito, 2024). Une fenêtre plus large permet au modèle d'intégrer un historique plus étendu, ce qui améliore la fluidité du discours, la stabilité du ton et la continuité thématique sur de longues séquences (NexusGPT, 2025d).

Cette capacité est particulièrement utile pour des tâches nécessitant une mémorisation à long terme, comme le suivi d'un raisonnement complexe, la compréhension d'une instruction divisée en plusieurs étapes ou la production de textes narratifs cohérents. Elle permet également au modèle de maintenir une certaine constance dans ses réponses, notamment dans le cadre d'interactions conversationnelles (Hipólito, 2024).

Toutefois, cette extension contextuelle implique une augmentation des ressources de calcul mobilisées. À l'inverse, une fenêtre plus restreinte offre une exécution plus rapide, mais peut nuire à la cohérence du texte en réduisant la capacité du modèle à se souvenir d'éléments évoqués précédemment (NexusGPT, 2025d).

7.4. Les différentes architectures de modèles LLM

7.4.1. Les modèles encodeurs

Les modèles encodeurs (ou autoencodeurs) utilisent un mécanisme d'attention pour produire une représentation globale du texte tenant compte du contexte. Leur entraînement repose sur la prédiction de mots masqués, ce qui les rend adaptés à des tâches comme l'extraction d'informations, la classification ou l'analyse d'opinion (Johnson & Hyland-Wood, 2025 ; Moradi et al., 2025).

Le modèle BERT, introduit en 2018, est un exemple emblématique. Il intègre à la fois des mots précédents et suivants pour prédire un mot masqué, et inclut une tâche complémentaire de relation logique entre phrases (Ramos et al., 2024).

Cependant, ces modèles sont peu adaptés à la génération de texte, car leur conception est centrée sur l’interprétation plutôt que la production (Moradi et al., 2025).

7.4.2. Les modèles décodeurs

Les modèles décodeurs (ou auto-régressifs) génèrent du texte de manière progressive à partir du contexte passé, sans anticiper les mots suivants (Moradi et al., 2025). Leur structure unidirectionnelle – gauche à droite – repose sur un mécanisme d’attention qui exclut toute information issue du futur (Ramos et al., 2024). Ils sont entraînés via le *Causal Language Modeling* (CLM), où chaque mot est prédit à partir d’une séquence incomplète. Ces modèles sont efficaces pour la génération, la traduction ou la synthèse (Johnson & Hyland-Wood, 2025).

Le modèle GPT, lancé en 2018, est entièrement basé sur des décodeurs, avec encodage positionnel des mots et attention unidirectionnelle (Ramos et al., 2024).

Malgré leurs performances, ils demandent beaucoup de ressources et peuvent perdre en cohérence sur des séquences longues (Moradi et al., 2025).

7.4.3. Les modèles encodeur-décodeur

Les modèles *seq2seq* (encodeur-décodeur) transforment une séquence d’entrée en une séquence de sortie en 2 étapes : l’encodeur analyse le texte source, le décodeur génère la réponse (Johnson & Hyland-Wood, 2025 ; Moradi et al., 2025).

BART incarne cette approche hybride : encodeur bidirectionnel (type BERT) et décodeur auto-régressif (type GPT) (Ramos et al., 2024). Leur complexité limite leur exploitation à grande échelle, ce qui pousse à privilégier les architectures spécialisées (Moradi et al., 2025).

7.4.4. Les modèles multitâches et multimodaux

Les modèles récents peuvent effectuer plusieurs tâches avec une seule architecture : résumer, répondre à des questions ou traduire, sans adaptation spécifique. Certains gèrent aussi des données autres que textuelles – images, sons, vidéos (Ramos et al., 2024).

Par exemple, Next-GPT traite plusieurs modalités via des modules spécialisés, convertissant chaque donnée dans un format exploitable par le modèle (Ramos et al., 2024).

Ces évolutions renforcent la capacité des agents IA à fonctionner de manière autonome dans des environnements complexes, en s’appuyant sur diverses sources d’information (Ramos et al., 2024).

7.5. Risques et limites

Les modèles de type LLM soulèvent plusieurs enjeux et vulnérabilités, notamment en ce qui concerne l’origine et la sensibilité des données mobilisées. Leur entraînement s’appuie sur

l’exploitation de vastes ensembles d’informations, dont certaines peuvent avoir été recueillies sans consentement explicite, ou contenir des éléments à caractère personnel ou confidentiel. Ces pratiques suscitent de vives inquiétudes quant au respect des droits liés à la vie privée (Chiat Ling Ong et al., 2024 ; Moradi et al., 2025).

Par ailleurs, les corpus employés peuvent intégrer des biais sociaux – stéréotypes, préjugés ou propos discriminatoires – susceptibles d’induire des résultats inéquitables. Un autre danger réside dans la génération de contenus erronés, pouvant être mobilisés pour désinformer ou influencer indûment l’opinion publique (Moradi et al., 2025).

Les LLM sont également vulnérables au phénomène d’hallucination, c’est-à-dire à la production de réponses plausibles en apparence mais erronées, voire incohérentes par rapport à l’instruction initiale. Ce type d’erreur pose un risque particulier dans les contextes nécessitant une information fiable, comme la production de résumés ou la réponse à des questions factuelles (Hipólito, 2024).

Un autre phénomène préoccupant est celui de l’oubli catastrophique (*catastrophic forgetting*), observé notamment lors du processus d’affinement (*fine-tuning*), qui consiste à spécialiser un modèle généraliste pour des tâches spécifiques. Cette phase d’apprentissage peut involontairement conduire à une perte de performance sur des compétences que le modèle maîtrisait auparavant (Johnson & Hyland-Wood, 2025 ; Moradi et al., 2025).

À cela s’ajoute le risque structurel d’effondrement du modèle (*model collapse*), qui survient lorsque les modèles sont réentraînés non plus à partir de données humaines, mais à partir de contenus générés par d’autres LLM. Cette pratique tend à appauvrir la diversité et la richesse linguistique, ce qui peut entraîner une dégradation progressive de la qualité des résultats (Johnson & Hyland-Wood, 2025).

Parmi les menaces d’ordre opérationnel figurent les attaques de type *jailbreak*, qui visent à contourner les garde-fous intégrés aux modèles afin d’accéder à des contenus prohibés, tels que des discours haineux ou des informations manipulées (Johnson & Hyland-Wood, 2025).

Enfin, il convient de souligner que ces modèles sont généralement entraînés sur des jeux de données antérieurs à leur mise en service. En conséquence, ils peuvent ne pas refléter les évolutions récentes ou les événements postérieurs à la date d’entraînement (NexusGPT, 2025d).

8. Generative Pre-trained Transformer (GPT)

8.1. Historique et conception des modèles

Les modèles *Generative Pre-trained Transformer* (GPT), développés par OpenAI, représentent une avancée majeure dans le domaine de l’intelligence artificielle générative. Basés sur une architecture de transformeur décodeur, ces modèles sont pré-entraînés sur de vastes quantités

textuelles, ce qui leur permet d'apprendre la structure du langage de manière autonome (Idowu et al., 2024).

Le premier modèle, GPT-1 (2018), a jeté les bases de cette approche avec 117 millions de paramètres. GPT-2 (2019), plus puissant (1,5 milliard de paramètres), a permis une génération de texte plus fluide et cohérente (Briouya et al., 2024). Mais c'est GPT-3 (2020), avec ses 175 milliards de paramètres, qui a marqué un tournant dans la performance et la polyvalence des modèles de langage. Il introduit notamment les capacités de *zero-shot*, *one-shot* et *few-shot learning*, qui permettent d'exécuter des tâches avec peu ou pas d'exemples d'entraînement (Briouya et al., 2024 ; Brown et al., 2020).

Pour surmonter certaines limites observées dans GPT-3 (notamment les difficultés d'interprétation d'instructions complexes), GPT-3.5 a été développé (Kalyan, 2023). Il intègre un apprentissage par renforcement à partir de retours humains (RLHF), ce qui améliore significativement la pertinence et la cohérence des textes générés. C'est sur cette version qu'est basé le modèle ChatGPT (OpenAI, n.d.-b).

L'arrivée de GPT-4 en 2023 marque une nouvelle étape, avec un modèle multimodal capable de traiter non seulement du texte, mais aussi des images. Il se distingue également par une compréhension plus fine du contexte et une meilleure stabilité dans la génération (Briouya et al., 2024).

Parmi les versions récentes, GPT-4o est aujourd'hui largement utilisé. Ce modèle prend en charge les entrées textuelles et visuelles, et génère des réponses sous forme de texte. Il est adapté à une grande variété de cas d'usage. GPT-4.1, destiné à des tâches plus complexes, est complété par des versions mini et nano, qui offrent respectivement un bon compromis entre performance et coût, ou une exécution rapide à faible coût (OpenAI, n.d.-a).

Enfin, le modèle GPT-4.5 se distingue par une amélioration notable des interactions conversationnelles. Grâce à un entraînement renforcé, il est capable de mieux interpréter l'intention de l'utilisateur et s'avère particulièrement efficace pour la rédaction, le codage ou encore la résolution de problèmes pratiques (OpenAI, 2025).

8.2. Limites et risques des modèles GPT

À l'instar des autres modèles de type LLM (Large Language Models), les modèles GPT présentent certaines limites et risques dans leur utilisation. Idowu et al. (2024) en identifient plusieurs.

Premièrement, au-delà d'environ 4 000 tokens, les segments initiaux du texte sont ignorés, ce qui nuit à la cohérence des échanges prolongés. Cette limite, appelée fenêtre contextuelle, correspond à la capacité maximale des modèles GPT à traiter des séquences continues (Idowu et al., 2024), comme exposé précédemment.

L'entraînement des modèles GPT repose sur des corpus issus de sources multiples, dont la composition influence directement la qualité et la neutralité des contenus générés. En l'absence de mécanismes de filtrage exhaustifs, ces données peuvent véhiculer des représentations biaisées ou stéréotypées. Par ailleurs, en raison de leur conception reposant sur des données statiques, ces modèles ne bénéficient pas d'une actualisation continue. Cette caractéristique limite leur capacité à intégrer des évolutions récentes ou à répondre de manière pertinente à des événements contemporains (Idowu et al., 2024).

Les modèles GPT sont également sujets à des phénomènes dits d'*hallucination* : ils peuvent produire des réponses qui, bien que cohérentes sur la forme, contiennent des informations inexactes, voire entièrement inventées (Idowu et al., 2024).

Par ailleurs, l'entraînement de ces modèles requiert des ressources informatiques particulièrement intensives, mobilisant sur de longues durées des infrastructures de calcul à haute performance. Cette exigence engendre non seulement des coûts économiques élevés, liés à la maintenance et à l'exploitation des équipements, mais soulève également des préoccupations environnementales en raison de la forte consommation énergétique associée à ces processus (Idowu et al., 2024).

Enfin, l'ampleur et la diversité des données employées lors de l'entraînement soulèvent des questions cruciales en matière de respect de la vie privée. L'absence de dispositifs robustes d'anonymisation ou de contrôle peut conduire à la reproduction accidentelle de données personnelles, posant ainsi un défi majeur pour la confidentialité des utilisateurs (Idowu et al., 2024).

9. Prompt engineering

9.1. Concept

Afin d'optimiser les performances des modèles de langage de grande taille (LLM), il est essentiel de maîtriser l'art de concevoir des prompts efficaces, une pratique connue sous le nom de *prompt engineering*. Cette approche consiste à formuler soigneusement des requêtes afin d'orienter les réponses du modèle vers les résultats attendus. L'objectif est d'obtenir des réponses adaptées au contexte et présentant une grande fiabilité. Parmi les stratégies utilisées figurent l'emploi de termes spécifiques qui déclenchent des réponses précises, la structuration réfléchie des interrogations posées au modèle, ainsi que la mise en contexte explicite des requêtes (Soliman et al., 2025).

9.2. Conseil de rédaction de prompt

La plateforme NexusGPT (2025d), dans son guide de formation, présente plusieurs recommandations essentielles pour la conception de prompts efficaces. Premièrement, il

convient de structurer la requête avec suffisamment de clarté et de détail pour prévenir toute réponse imprécise ou hors-sujet. La pertinence des résultats obtenus dépend largement du niveau de détail contextuel fourni au modèle. Intégrer des éléments explicatifs précis, des exemples concrets, ou des indications sur la structure attendue évite que l'IA se repose sur des suppositions non vérifiées, minimisant ainsi le risque qu'elle produise des informations erronées (NexusGPT, 2025d).

Par ailleurs, une stratégie efficace pour réduire les risques d'informations inventées consiste à demander clairement au modèle de préciser lorsqu'une réponse ne relève pas de ses connaissances. De plus, il est utile de spécifier précisément le format désiré, tel qu'une présentation sous forme de liste ou un nombre précis de mots. Dans le cas de demandes complexes ou comportant plusieurs volets, fractionner la tâche en différentes étapes distinctes permet au modèle d'améliorer sa précision et d'offrir un résultat mieux organisé (NexusGPT, 2025d).

Enfin, il est possible de favoriser l'autocritique du modèle en l'invitant à vérifier par lui-même la fiabilité ou la logique de sa réponse initiale. Étant donné qu'il est inhabituel d'obtenir immédiatement une réponse optimale, NexusGPT suggère une approche progressive consistant à reformuler et à ajuster le prompt après chaque essai, jusqu'à obtenir une réponse satisfaisante (NexusGPT, 2025d).

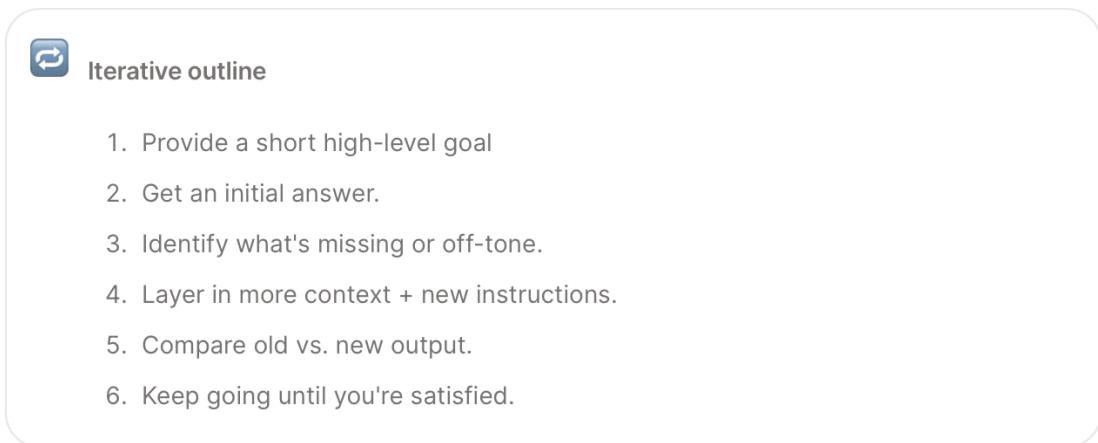


Figure 6 - Schéma d'une approche progressive de la réalisation d'un prompt.

Source : NexusGPT (2025d), extrait de *Foundational Knowledge*.

9.3. Techniques de prompting

Il existe plusieurs techniques de conception de prompts, particulièrement utiles pour les requêtes complexes. Soliman et al. (2025) et NexusGPT (2025d) identifient notamment le *zero-shot prompting*, le *few-shot prompting*, la méthode *chain-of-thought* et les *cadres structurés*.

Le *zero-shot* prompting consiste à formuler une requête directe, sans fournir d'exemple préalable. À l'inverse le *few-shot* prompting s'appuie sur un petit nombre d'exemples pour guider le modèle et renforcer la pertinence de la réponse (NexusGPT, 2025d ; Soliman et al., 2025).

La méthode *chain-of-thought* est particulièrement adaptée aux tâches nécessitant un raisonnement logique ou des calculs. Elle consiste à guider le modèle en décomposant la requête en étapes successives, facilitant une progression raisonnée vers la réponse attendue (NexusGPT, 2025d).

Enfin, les *cadres de prompts structurés* permettent de standardiser la formulation des requêtes, notamment dans les contextes collaboratifs ou pour un usage à grande échelle. Cette approche repose sur une segmentation du prompt en sections distinctes, comme le rôle, la tâche, le contexte ou les exemples, afin de clarifier les attentes et optimiser la réponse du modèle (NexusGPT, 2025d).

9.4. Considérations éthiques dans la conception de prompts

Dans son guide, la plateforme NexusGPT (2025d) rappelle que la conception des prompts ne vise pas uniquement à obtenir des réponses pertinentes, mais doit également garantir que celles-ci soient fiables, respectueuses et précises.

— Limiter les biais

Comme évoqué précédemment, les données d'entraînement des modèles peuvent contenir des biais implicites. Afin d'en limiter les effets, il est essentiel de formuler les requêtes de manière impartiale. Lorsque le contenu généré présente un biais identifiable, il peut être utile de reformuler la demande ou d'orienter le modèle vers une diversité d'interprétations (NexusGPT, 2025d).

— Respect des règles et des limites

Certains prompts peuvent involontairement conduire à la génération de contenus sensibles, contraires à l'éthique ou même illégaux. Il est donc primordial de préserver la confidentialité des données, notamment lorsqu'elles sont sensibles. En cas de réponse inappropriée, la correction du prompt et l'ajustement des consignes sont essentiels (NexusGPT, 2025d).

— Vigilance face aux informations erronées

Le phénomène d'hallucination déjà évoqué peut conduire le modèle à produire des informations inexactes. Il est donc crucial de vérifier les réponses générées, soit par une validation humaine, soit en demandant explicitement des sources ou références (NexusGPT, 2025d).

— Transparency et usage responsable

Lorsqu'un contenu généré par IA est diffusé, il est important d'en indiquer la provenance. Par ailleurs, dans les domaines sensibles comme la santé, le droit, la finance... ces productions doivent être considérées comme des outils de support, et non comme des avis professionnels définitifs (NexusGPT, 2025d).

10. Les limites de l'intelligence artificielle

Bien que l'intelligence artificielle ait révolutionné de nombreux secteurs de notre société, contribuant à des avancées technologiques et scientifiques majeures, elle demeure une technologie imparfaite, sujette à diverses limites et à des remises en question. Des auteurs tels que Gupta (2025) ainsi que Burke et Akhtar (2023) en ont identifié plusieurs.

— Dépendance à de larges quantités de données

Pour atteindre un niveau de performance élevé, les modèles d'intelligence artificielle doivent être entraînés sur de très grandes quantités de données. Cependant, les résultats obtenus dépendent fortement de la qualité de ces données. Or, rassembler un volume important de données fiables et représentatives constitue souvent un véritable défi (Gupta, 2025).

— Susceptibilité aux biais

Les modèles d'intelligence artificielle dépendent fortement de vastes ensembles de données pour leur entraînement. Or, ces bases de données peuvent contenir des biais. Ces derniers se traduisent parfois par des résultats empreints de stéréotypes, de discriminations ou d'orientations idéologiques, reflétant ainsi certaines problématiques sociétales sensibles (Burke & Akhtar, 2023 ; Gupta, 2025).

— Limites des modèles face aux documents longs

Malgré l'élargissement progressif des fenêtres contextuelles, les modèles d'intelligence artificielle, en particulier les grands modèles de langage (LLM), rencontrent des difficultés persistantes lorsqu'ils doivent traiter des documents longs. Plusieurs recherches comme celle de Levy et al., 2024 ; Liu et al., 2023 ; Zhou et al., 2025 mettent en lumière cette faiblesse.

Liu et al. (2023) parlent d'un effet de « courbe en U », indiquant que les modèles sont performants lorsque l'information à chercher se situe au début ou à la fin du document, mais qu'ils peinent lorsque celle-ci se trouve au centre.

Levy et al. (2024) confirment que l'augmentation de la longueur d'un texte, même sans modification de la tâche, peut nuire aux performances de raisonnement. Ils soulignent que les modèles ont du mal à localiser les informations pertinentes ou à suivre correctement les instructions.

Enfin, Zhou et al. (2025) montrent qu'au-delà d'une certaine longueur, les représentations sémantiques internes s'effondrent, ce qui altère la capacité du modèle à distinguer les idées et affecte négativement les résultats obtenus.

Ces constats remettent en question l'idée selon laquelle l'extension de la fenêtre contextuelle suffirait à garantir une meilleure compréhension des textes longs.

— Sécurité, confidentialité et éthique

L'essor de l'intelligence artificielle soulève de nombreuses interrogations liées au respect de la vie privée, à la sécurité des données et aux enjeux éthiques. Dans le domaine du marketing, par exemple, ces technologies sont capables de collecter et d'analyser des données personnelles afin de générer des recommandations personnalisées. Cela suscite des préoccupations croissantes quant à la confidentialité des informations traitées par l'IA (Burke & Akhtar, 2023).

Par ailleurs, les systèmes d'intelligence artificielle peuvent être exposés à des attaques provenant d'acteurs mal intentionnés, qui manipulent les données d'entrée dans le but de tromper les algorithmes. Ce type de vulnérabilité peut avoir des conséquences graves dans des secteurs sensibles tels que la finance ou la défense (Gupta, 2025).

Afin de réguler l'utilisation de ces technologies, plusieurs cadres normatifs et recommandations ont été mis en place, notamment les principes éthiques élaborés par des organismes comme l'IEEE ou l'OCDE. Néanmoins, leur application effective au sein des systèmes d'intelligence artificielle demeure complexe et souvent difficile à concrétiser (Gupta, 2025).

Cette question de la sécurité et de la confidentialité des données apparaît comme un point de vigilance majeur pour l'ensemble des responsables que j'ai pu interviewer. Ainsi, Mukile (2025) explique que l'évaluation de la sécurité a été prioritaire lors de l'adoption de l'outil d'intelligence artificielle Karomia :

« Pour l'instant, les seules préoccupations qu'on a vraiment étudiées, c'est tout ce qui touche à la sécurité des systèmes et à la confidentialité des données. On a fait des vérifications avec Karomia pour s'assurer que ça respectait bien nos standards en gestion de l'information et en protection des données. »

Le groupe T&S manifeste également une forte sensibilité à la confidentialité, au point de restreindre volontairement l'usage de l'IA à certains types de données. Mangold (2025) indique :

« (...) on parle de données confidentielles. Autant, il y en a qui ne le sont pas du tout. Tout ce qui est lié à l'environnement, c'est publié sur notre site, ça va être republié sur notre site cette année, etc. Donc, il n'y a pas de question de confidentialité sur ce volet-là. Mais pour tout ce qui est lié à la partie sociale, oui, en partie, il y a une question de confidentialité. C'est-à-dire qu'il y a certaines données qu'on ne va pas transmettre. »

Chez Alstom, l'attention portée à la protection des données se traduit par un environnement sécurisé dans lequel les outils IA sont directement intégrés aux systèmes internes de l'entreprise. Daga (2025) précise :

« In terms of transparency and confidentiality of the data, we have our cybersecurity systems in place. For example, with Copilot, it's already integrated into our servers, our systems. So, it's not something that our data is leaving our service to someone else's, you know. So, in that way, we try to keep our data protected in that manner. And plus, there are regulations on, you know, data privacy, data management, I think GDPR, that's what it's called, that protects us from, you know, losing data or making sure that we're not violating cybersecurity rules. So, I think in that way, we are quite okay with how the systems are right now. »

Enfin, chez Veolia, l'usage de ChatGPT est strictement encadré par le service informatique, qui en a développé une version interne :

« Il faut savoir que chez nous, on utilise en fait ChatGPT. On l'utilise dans une version contrôlée par notre département informatique, qui l'a personnalisée et fermée au niveau de l'utilisation pour Veolia. Donc là, oui, la confidentialité des données est clé, et c'est quelque chose sur lequel on est très stricts sur ce point » (Jamin, 2025).

Lineas s'exprime également sur le choix de sa plateforme d'intelligence artificielle, en mettant l'accent sur les garanties en matière de sécurité :

« Nous utilisons la plateforme Karomia. C'est un environnement sécurisé. Nous ne mettons pas tous nos documents sur Internet ni sur un pilote IA ouvert. C'est un environnement séparé. La confidentialité est garantie » (Vanoppen, 2025).

— L'environnement

L'impact environnemental de l'intelligence artificielle suscite de plus en plus d'attention, notamment en raison de la consommation énergétique élevée requise pour l'entraînement des modèles. Cette consommation entraîne des émissions importantes de gaz à effet de serre (GES) (Burke & Akhtar, 2023).

D'après ces auteurs, l'entraînement du modèle BERT dans un centre de données situé aux États-Unis a généré environ 22 à 28 kilogrammes de dioxyde de carbone, soit presque deux fois les émissions produites pour un entraînement équivalent réalisé en Norvège. Ils précisent également qu'un modèle de traitement automatique du langage pourrait, à lui seul, émettre une quantité de gaz à effet de serre comparable à celle produite par 5 voitures sur l'ensemble de leur durée de vie (Burke & Akhtar, 2023).

Mukile (2025) s'est également interrogé sur les conséquences environnementales de l'usage de l'intelligence artificielle. Selon lui :

« *Si on est dans une logique de durabilité, on doit aussi se poser la question de l'empreinte écologique de l'outil qu'on utilise. Idéalement, il faudrait pouvoir développer une IA avec un très faible impact en matière d'énergie.* »

— **Le besoin de relecture humaine du contenu généré par l'IA**

L'un des points les plus fréquemment mentionnés au cours de mes entretiens est la nécessité d'un contrôle humain sur les contenus produits par l'intelligence artificielle. En dépit des gains d'efficacité et de rapidité qu'elle offre, la confiance envers l'IA n'est pas totale, et les personnes interrogées soulignent qu'une relecture humaine reste nécessaire pour garantir la cohérence, la pertinence et l'exactitude des informations générées.

Vanoppen (2025) insiste à ce sujet :

« *Of course. You need to read through it. Because AI is beautiful. It's magnificent. But you need to read through it. It is not always making sense. And so, the checks and verifications, you still need to do it.* »

Ce constat est partagé par chez Sibelga, à propos de l'outil Karomia :

« *Mais dans tous les cas, il faudra une relecture humaine pour s'assurer de la cohérence et éviter qu'on raconte n'importe quoi* » (Mukile, 2025).

Après avoir exploré les fondements, les capacités et les limites de l'intelligence artificielle et de ses modèles, il convient désormais d'examiner une forme d'implémentation concrète de ces technologies : les agents IA. Ces derniers constituent des cas d'usage développés au cours de mon stage et illustrent la mise en œuvre opérationnelle des principes abordés précédemment. Le chapitre suivant s'intéressera donc à la notion d'agent intelligent, à ses caractéristiques principales, ainsi qu'aux enjeux liés à leur conception et leur utilisation.

Chapitre 3 : Les agents d'intelligence artificielle

1. Définitions et origines des agents d'IA

La notion d'agent trouve ses fondements dans la tradition philosophique, où elle renvoie à une entité dotée de la faculté d'agir intentionnellement, selon sa propre volonté. Des penseurs tels qu'Aristote ou Hume ont contribué à façonner cette conception, centrée sur l'autonomie de décision (Ramos et al., 2024 ; Xi et al., 2025). Dans le champ de l'intelligence artificielle, cette idée a été reprise et structurée notamment par Russell dans les années 1990 (Zhao et al., 2023).

Cette capacité à percevoir, planifier et agir de manière adaptée caractérise aujourd'hui les agents artificiels. Lorsqu'ils reposent sur des modèles de langage de grande taille (LLM), on parle alors d'agents linguistiques. Ces derniers traitent des observations issues d'un environnement spécifique ; défini par des outils ou des objectifs à atteindre, afin de produire des actions en retour (Ramos et al., 2024).

Faute de définition unique, certains chercheurs considèrent que l'agentivité ne devrait pas être envisagée de manière absolue : un système ne serait pas simplement agent ou non-agent, mais plutôt appréhendé comme une notion graduelle, variant selon son degré d'autonomie ou de capacité d'action (Kapoor et al., 2024).

Cette distinction progressive se retrouve dans les travaux de Wooldridge et Jennings (1995), qui identifient 2 approches complémentaires du concept d'agent.

La première, qualifiée de conception faible, repose sur une définition fonctionnelle. L'agent y est envisagé comme un programme autonome, généralement logiciel, capable de traiter des informations issues de son environnement, d'y réagir en temps réel et d'accomplir des tâches ciblées, seul ou en interaction avec d'autres systèmes (Wooldridge & Jennings, 1995).

La seconde, plus ambitieuse, mobilise des notions inspirées des sciences cognitives. Certains agents y sont conçus comme des entités disposant d'états internes comparables à des croyances, des désirs ou des intentions, et capables d'agir selon des mécanismes s'apparentant à une forme de raisonnement. Cette perspective inclut parfois des attributs supplémentaires, tels que la mobilité sur un réseau, la simulation d'émotions ou la coopération avec d'autres agents dans un environnement partagé (Wooldridge & Jennings, 1995).

L'intelligence artificielle dite *agentique* désigne des systèmes à forte autonomie, capables de poursuivre des objectifs élaborés sans recourir constamment à une supervision humaine. Les avancées en matière de modèles de langage leur permettent d'analyser des situations, de prendre des décisions adaptées et de conduire des actions alignées sur leurs finalités (Horne, 2025).

Dans ce contexte, une nouvelle génération d'agents a émergé : les agents génératifs. Ceux-ci associent des capacités de production de contenu à des fonctions cognitives avancées, telles que

la planification, l'accès à des ressources numériques, l'adaptation en temps réel et la réalisation d'actions coordonnées. Contrairement aux modèles classiques, ces agents peuvent structurer des tâches complexes par eux-mêmes, réagir aux informations reçues en cours d'exécution et perfectionner leurs réponses via des boucles d'apprentissage ou des mécanismes d'ajustement fondés sur l'analyse de leurs propres performances. Des implémentations notables comme AutoGPT, BabyAGI ou AgentGPT illustrent ces évolutions (Hettiarachchi, 2025).

2. La typologie des agents d'IA

2.1. Les agents symboliques

Aux débuts de la recherche en intelligence artificielle, l'approche symbolique s'est imposée comme paradigme dominant. Elle se fondait sur des systèmes logiques formels, utilisant des représentations symboliques et des règles déductives pour structurer et exploiter des connaissances. C'est dans ce cadre qu'ont été développés les premiers agents, conçus pour reproduire certaines dimensions du raisonnement humain. Ils se distinguaient par leur capacité à exprimer des connaissances de manière structurée et claire. Toutefois, ces systèmes ont rapidement révélé leurs limites face à des environnements complexes ou incertains. Leur traitement algorithmique, souvent lourd et rigide, peinait à produire des résultats opérationnels dans des temps compatibles avec les exigences du monde réel (Xi et al., 2025).

2.2. Les agents réactifs

Une autre catégorie d'agents, apparue progressivement dans le développement de l'intelligence artificielle, est celle des agents dits « réactifs ». Ces agents privilégient une interaction directe avec leur environnement. Ils fonctionnent selon un principe de perception-action, qui leur permet de répondre rapidement aux stimuli externes. Cette conception les rend particulièrement efficaces pour des tâches simples ou nécessitant des réactions immédiates. En revanche, leur architecture ne leur permet pas de gérer des situations impliquant de la planification ou un raisonnement de haut niveau (Xi et al., 2025).

2.3. Les agents basés sur l'apprentissage par renforcement

L'essor des capacités de calcul et l'accès élargi aux données ont favorisé l'apparition de nouvelles méthodes pour concevoir des agents intelligents. L'une des approches les plus marquantes est celle de l'apprentissage par renforcement, qui permet de former des agents à partir de leurs interactions successives avec l'environnement. Cette méthode vise à maximiser, au fil du temps, les bénéfices retirés de leurs actions. Grâce à cette dynamique d'apprentissage

expérientiel, les agents peuvent évoluer dans des contextes inconnus, sans nécessiter de supervision humaine constante. Ces agents trouvent aujourd’hui des applications dans des domaines variés, tels que le jeu, la robotique ou la simulation. Cependant, l’apprentissage par renforcement soulève encore plusieurs défis. Les phases d’entraînement sont souvent longues, les performances peuvent varier selon la qualité des données et des difficultés de stabilité apparaissent, notamment lors de la transposition de ces systèmes dans des environnements concrets aux contraintes multiples (Xi et al., 2025).

2.4. Les agents avec apprentissage par transfert et métapprentissage

Afin de surmonter les limites de l’apprentissage par renforcement, 2 méthodes complémentaires ont émergé : l’apprentissage par transfert et le métapprentissage. La première vise à réutiliser les connaissances acquises lors de tâches antérieures pour accélérer l’apprentissage dans de nouveaux contextes. Elle facilite ainsi la transition d’un domaine à un autre en réduisant l’effort d’adaptation requis. Le métapprentissage, quant à lui, repose sur l’idée que l’agent peut apprendre à ajuster sa propre manière d’apprendre. Grâce à cette capacité, il peut élaborer rapidement des stratégies efficaces à partir de peu d’exemples, en s’appuyant sur des schémas généraux déjà assimilés (Xi et al., 2025).

Néanmoins, ces approches présentent encore des défis. L’apprentissage par transfert peut échouer, voire nuire aux performances, si les tâches d’origine et les cibles sont trop éloignées. Le métapprentissage, de son côté, nécessite souvent un pré-entraînement long et de vastes ensembles de données, ce qui peut limiter sa portabilité (Xi et al., 2025).

2.5. Les agents basés sur des modèles de langage de grande taille (LLM)

Les modèles de langage de grande taille constituent aujourd’hui un socle central pour le développement d’agents capables d’interpréter des instructions en langage naturel tout en agissant sur leur environnement. Ces agents combinent des mécanismes de raisonnement inspirés des approches symboliques avec des interactions adaptatives proches du modèle réactif. En mobilisant l’apprentissage par transfert et des interfaces multimodales, ils peuvent s’adapter à des tâches nouvelles avec peu de données, voire sans ajustement explicite. Leur usage s’étend désormais à des secteurs concrets comme la programmation ou la recherche scientifique (Xi et al., 2025).

D'autres typologies ont également été proposées dans la littérature récente, fondées sur des critères fonctionnels ou structurels. Krishnan (2025) distingue notamment :

- des agents réflexes simples, qui fonctionnent selon des règles fixes « si-alors » et sont adaptés à des environnements simples et entièrement observables ;
- des agents orientés vers des objectifs ;
- et des agents utilitaristes, qui évaluent plusieurs options et choisissent celle qui maximise le résultat le plus favorable pour leurs objectifs.

Ces classifications, qui complètent celles présentées ici, rappellent que les typologies d'agents restent évolutives et s'adaptent aux avancées récentes en matière d'architectures et de cas d'usage, comme le souligne l'auteur.

3. Architecture d'un agent IA

Dans la littérature récente, plusieurs travaux ont cherché à formaliser l'architecture des agents, notamment ceux reposant sur des modèles de langage de grande taille (LLM).

Wang et al. identifient 4 modules fonctionnels structurants dans ces systèmes : le profilage, le stockage d'information, l'élaboration de stratégie et la mise en œuvre des décisions. Weng propose une variante en remplaçant le profilage par l'intégration d'outils interactifs (cités dans Ramos et al., 2024).

Ramos et al. (2024) décrivent quant à eux une organisation fondée sur un noyau décisionnel, souvent un LLM, entouré de sous-systèmes spécialisés qui permettent à l'agent de traiter des informations, de raisonner, d'élaborer des stratégies et d'interagir avec son environnement. L'agent fonctionne selon un cycle continu qui lui permet d'analyser des signaux, de choisir une action appropriée et de réagir en conséquence. Il combine pour cela des outils internes, dédiés au traitement de l'information, et des outils externes, destinés à une mise en œuvre concrète dans le monde environnant.

Xi et al. (2025) présentent une architecture conceptuelle fondée sur 3 modules interdépendants : la perception, le centre décisionnel – désigné comme « cerveau » – et l'action. Le bloc central, le cerveau, repose sur un grand modèle de langage (LLM), qui joue un rôle déterminant dans le traitement de l'information, l'élaboration de raisonnements, la prise de décisions, et la gestion des connaissances stockées. L'unité de perception vise à étendre les capacités de l'agent en lui permettant de capter et d'interpréter des données provenant de diverses modalités (texte, image, son). Ces informations sont ensuite traduites dans un format compréhensible par le modèle. Enfin, la fonction d'action donne à l'agent la capacité d'intervenir sur son environnement par le biais d'outils ou de dispositifs d'exécution. L'ensemble fonctionne selon un cycle continu d'observation, de traitement et d'action, permettant une interaction dynamique avec le milieu environnant.

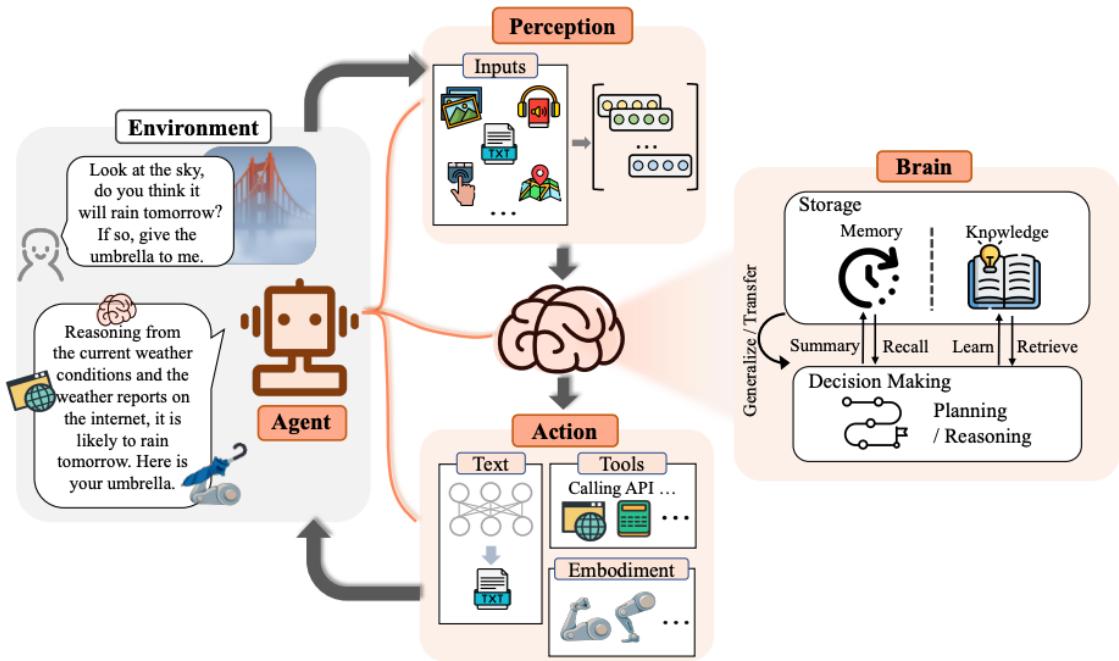


Figure 7 - Cadre conceptuel d'un agent basé sur un modèle LLM

Source : Xi et al., (2025), extrait de *The rise and potential of large language model based agents : A survey*

Hettiarachchi (2025) décrit une architecture type pour les agents génératifs, articulée autour d'un grand modèle de langage (LLM) qui occupe une position centrale en tant que moteur d'interprétation et de gestion des interactions. Ce cœur est complété par plusieurs composants spécialisés. La mémoire conserve à la fois des informations utiles à court terme et des données structurantes pour les interactions futures. La planification, quant à elle, organise les missions complexes en séquences d'étapes adaptables. Enfin, les interfaces assurent l'exécution concrète des décisions dans des systèmes extérieurs aux caractéristiques variables.

En somme, bien que les architectures proposées varient selon les auteurs, elles reposent toutes sur des fondements fonctionnels similaires – perception, mémoire, raisonnement, planification et action – généralement articulés autour d'un modèle central de type LLM.

4. Les LLM : la base des agents intelligents

Quelle que soit l'architecture envisagée dans le point précédent, un élément commun se dégage : le rôle structurant joué par le modèle de langage de grande taille (LLM) au sein des agents d'intelligence artificielle. Mais qu'est-ce qui justifie cette centralité ?

Xi et al. (2025) proposent de l'expliquer à travers 4 propriétés fondamentales :

La première est l'autonomie, entendue comme la capacité d'agir de manière indépendante, sans dépendre d'instructions constantes de la part d'un humain. Les LLM manifestent cette caractéristique en générant des contenus structurés et originaux, en traitant des tâches définies par des consignes globales, sans étapes détaillées, et en adaptant leurs réponses en fonction des informations reçues (Xi et al., 2025).

La deuxième est la réactivité, c'est-à-dire l'aptitude d'un agent à détecter des variations dans son environnement et à adapter son comportement en conséquence. Grâce à l'intégration de modalités comme l'image ou le son, les LLM étendent aujourd'hui leur capacité d'interaction au-delà du seul langage écrit (Xi et al., 2025).

Vient ensuite la proactivité, qui désigne l'aptitude à initier des démarches intentionnelles visant à atteindre des résultats spécifiques. Lorsqu'ils sont guidés par des instructions adaptées, les LLM sont capables d'organiser leur raisonnement et de construire un plan d'action détaillé. Ils peuvent aussi décomposer un objectif complexe en étapes réalisables, puis ajuster leur approche en fonction des évolutions du contexte (Xi et al., 2025).

Enfin, la dimension sociale des agents s'exprime par leur aptitude à interpréter et à générer un langage naturel accessible, favorisant ainsi des échanges pertinents et collaboratifs tant avec d'autres agents qu'avec des humains. Cette compétence ouvre la voie à des interactions riches, basées sur la coopération et le dialogue (Xi et al., 2025).

5. Quelques applications concrètes des agents d'intelligence artificielle

Les agents d'intelligence artificielle, en particulier ceux reposant sur des modèles de langage de grande taille (LLM), se déploient désormais dans de nombreux secteurs professionnels. Hettiarachchi (2025) et Krishnan (2025) mettent en lumière ces avancées.

Dans le domaine entrepreneurial, l'intégration des agents d'intelligence artificielle générative a conduit à une transformation significative des pratiques organisationnelles, en particulier grâce à l'automatisation de nombreuses fonctions opérationnelles. Des systèmes tels qu'AutoGPT, AgentGPT ou Einstein GPT de Salesforce sont désormais utilisés pour rationaliser des tâches quotidiennes. Celles-ci vont de la gestion de la communication électronique à la coordination des données clients dans les CRM, en passant par l'établissement de prévisions financières et l'analyse stratégique des données (Hettiarachchi, 2025).

Dans les ressources humaines, les agents IA permettent d'optimiser l'accompagnement des collaborateurs à différentes étapes du parcours professionnel. Certaines entreprises technologiques exploitent des agents d'IA capables de gérer plus de 80 % des demandes courantes des employés. Cette automatisation permet non seulement d'accélérer les réponses, mais aussi de redéployer les équipes RH vers des missions à plus forte valeur ajoutée (Krishnan, 2025).

Dans le domaine de la relation client, l'opérateur Vodafone a mis en œuvre un système d'assistance basé sur l'intelligence artificielle, qui traite de manière autonome plus de 70 % des sollicitations. Cette solution a contribué à accélérer les délais de traitement de près de 50 %, sans altérer la qualité perçue du service par les clients (Krishnan, 2025).

Dans le champ de la recherche scientifique, les agents génératifs jouent un rôle d'accélérateur à plusieurs étapes du processus. Ces agents contribuent à accélérer les différentes étapes du travail scientifique, en soutenant la formulation d'hypothèses, l'analyse de la littérature existante et l'élaboration de plans expérimentaux. Leur capacité à extraire et structurer des informations à partir de corpus scientifiques volumineux constitue un appui précieux pour les chercheurs, en particulier dans l'identification de lacunes de recherche ou de pistes d'exploration inédites (Hettiarachchi, 2025).

Dans le domaine de la santé, les agents assistent les professionnels dans l'exécution de tâches administratives et cliniques. Ils participent à la collecte de données avant consultation, à l'orientation des patients, ou encore à l'analyse de symptômes. Des solutions comme Med-PaLM, DAX Copilot ou Babylon Health illustrent ces usages en combinant traitement du langage naturel, reconnaissance d'images médicales et suivi en temps réel grâce aux capteurs connectés (Hettiarachchi, 2025).

Le domaine juridique profite également des avancées de l'IA, avec des agents capables de réaliser des recherches jurisprudentielles ciblées, d'analyser des contrats et d'identifier des clauses à risque. Ils facilitent également la rédaction de documents juridiques et l'évaluation de la conformité réglementaire, réduisant ainsi les délais de traitement et les risques d'erreur (Krishnan, 2025).

Les secteurs créatifs et éducatifs ne sont pas en reste. Dans les industries culturelles, les agents IA contribuent à la génération de contenus originaux en musique, en graphisme ou en conception de jeux vidéo, tout en assurant la cohérence stylistique des productions. En éducation, ils permettent la mise en place d'environnements d'apprentissage personnalisés, adaptés au profil de chaque apprenant. L'initiative Khanmigo, développée par la Khan Academy, illustre cette dynamique, en offrant un accompagnement individualisé grâce à l'interaction continue entre l'élève et l'agent (Hettiarachchi, 2025).

L'agriculture de précision et la gestion urbaine exploitent également les capacités des agents, notamment pour surveiller l'état des sols, prédire les rendements ou optimiser la consommation d'énergie (Krishnan, 2025).

Ces quelques exemples illustrent l'étendue des domaines dans lesquels les agents d'intelligence artificielle transforment les pratiques professionnelles.

6. Limites et défis des agents d'intelligence artificielle.

6.1. Limites techniques des agents

Les agents d'intelligence artificielle présentent également certaines limites, parmi lesquelles figure le phénomène d'hallucination. Comme évoqué dans plusieurs sections de ce mémoire, ce phénomène désigne la production, par l'agent, d'informations erronées ou entièrement inventées, souvent formulées de manière crédible. Ce caractère vraisemblable est susceptible d'induire les utilisateurs en erreur, en particulier lorsque ceux-ci s'appuient sur ces systèmes pour prendre des décisions importantes (Hettiarachchi, 2025).

Les agents reposant sur des modèles de type LLM s'appuient sur des instructions formulées en langage naturel (prompt) pour définir leurs objectifs et réaliser les tâches assignées. Toutefois, en cas de formulations ambiguës, d'un manque de contexte ou d'une structure peu claire, l'agent peut en faire une interprétation erronée. Cette incompréhension affecte alors la pertinence de la réponse et peut compromettre le bon déroulement de la tâche demandée (Hettiarachchi, 2025).

On retrouve également ici le problème déjà évoqué précédemment concernant les biais présents dans les données d'entraînement, pouvant conduire à des réponses marquées par des formes de discrimination ou de préjugés sociaux (Hettiarachchi, 2025).

Krishnan (2025) met également en lumière certaines limites techniques des agents, notamment leur difficulté à gérer des raisonnements complexes. Ceux-ci incluent l'incapacité à traiter efficacement des enchaînements logiques, à identifier des liens de causalité ou à résoudre des problèmes abstraits comme les calculs mathématiques.

Une autre difficulté réside dans la capacité des agents à suivre et à structurer les informations au fil de longues interactions. Ils peuvent omettre des éléments importants ou produire des réponses incohérentes en raison d'une mémoire contextuelle limitée (Krishnan, 2025).

Enfin, leur interaction avec des outils externes reste imparfaite : ils peinent parfois à choisir l'outil le plus adapté à une tâche donnée ou à interpréter correctement les résultats, ce qui peut nécessiter une supervision humaine accrue (Krishnan, 2025).

Les agents génératifs actuels présentent également une limite importante : leur incapacité à rendre compte de leur processus de raisonnement et de prise de décision. Ce manque de transparence se traduit par l'impossibilité pour l'utilisateur de comprendre les choix opérés par l'agent, en raison de mécanismes internes complexes et difficilement accessibles, souvent assimilés à une « boîte noire » (Hettiarachchi, 2025).

6.2. Risques juridiques, éthiques et impacts sociaux

L'utilisation des agents génératifs soulève également des préoccupations juridiques et éthiques, notamment en ce qui concerne la sécurité des données. Ce risque se matérialise notamment lorsque ces systèmes sont entraînés à partir de données appartenant à des organisations : si ces données sont protégées par des droits de propriété intellectuelle, leur utilisation non autorisée peut constituer une violation. En outre, dans certaines juridictions dotées de réglementations strictes en matière d'intelligence artificielle – comme l'Union européenne avec le RGPD et la loi sur l'IA – les agents génératifs doivent respecter des exigences précises en matière de transparence, de sécurité et de responsabilité. De plus, l'interaction continue entre les utilisateurs et ces agents suscite des questions sur le respect de la vie privée, notamment en ce qui concerne le recueil du consentement et les modalités de traitement des données personnelles (Hettiarachchi, 2025).

Comme le souligne Krishnan (2025), l'usage croissant d'agents intelligents dans la vie quotidienne peut engendrer une perte progressive d'autonomie. En effet, la délégation de tâches, notamment celles impliquant un jugement humain ou une responsabilité éthique, pourrait entraîner une réduction des capacités décisionnelles individuelles, voire un appauvrissement de certaines compétences professionnelles.

L'automatisation croissante soulève également des enjeux économiques : dans certains secteurs, elle pourrait conduire à une réduction du besoin de main-d'œuvre pour les tâches répétitives, posant la question de l'adaptation du marché du travail (Krishnan, 2025).

7. Gestion des données sensibles dans le cadre du projet

Dans le cadre du développement des agents d'intelligence artificielle mené durant mon stage, les questions de confidentialité et de sécurité des données ont été soulevées à plusieurs reprises, tant au sein de l'équipe que lors des entretiens avec les responsables, comme mentionné précédemment. Une interrogation centrale a émergé : *comment garantir la sécurité des données qui nous sont transmises, puis intégrées à nos agents ?*

La spécificité de la plateforme utilisée réside dans l'adoption de standards de sécurité particulièrement élevés, offrant ainsi aux futurs clients les garanties nécessaires en matière de protection des données. NexusGPT repose en effet sur différents standards de sécurité :

Reconnue à l'échelle internationale, la norme ISO/IEC 27001 atteste qu'une organisation a adopté une démarche rigoureuse pour assurer la protection de ses données. Elle s'appuie sur un système de management qui permet d'identifier, d'évaluer et de traiter les risques relatifs à la sécurité de l'information. Cette norme fournit un ensemble de principes et de pratiques pour encadrer le développement ou l'usage des systèmes d'IA dans une logique de responsabilité et

de transparence. Cette certification constitue donc un gage de fiabilité pour les partenaires comme pour les clients (ISO, 2022).

L'ISO/IEC 42001, est une norme internationale dédiée à la gouvernance des systèmes d'intelligence artificielle. Elle fournit un cadre structuré permettant aux organisations, qu'elles conçoivent ou utilisent des solutions basées sur l'IA, de piloter ces technologies de façon responsable. Cette norme propose une approche structurée permettant aux organisations d'identifier les enjeux liés à l'IA, d'intégrer des principes éthiques dans leurs processus et de trouver un équilibre entre les dynamiques d'innovation technologique et les exigences de gouvernance et de transparence (ISO, 2023).

Adopté à l'échelle de l'Union européenne, le Règlement Général sur la Protection des Données (RGPD) a pour objectif de garantir aux citoyens un meilleur contrôle sur l'utilisation de leurs données personnelles. Il impose aux entités publiques et privées des obligations rigoureuses en matière de traitement, de sécurité et de transparence, tout en prévoyant des mécanismes indépendants de contrôle et de sanction. Ce cadre juridique harmonisé vise à instaurer un climat de confiance entre les organisations et les individus, et à prévenir les dérives liées à une utilisation abusive des données (EUR-Lex, 2022). Ces obligations comprennent notamment la nécessité de recueillir un consentement libre et éclairé (art. 6-7), de ne traiter les données que dans un but déterminé et légitime (art. 5 § 1.b), de garantir la sécurité des informations personnelles (art. 32), ainsi que de respecter les droits des individus à l'accès, la rectification ou l'effacement de leurs données (art. 15 à 17). En cas de non-conformité, les entreprises peuvent risquer des pénalités pouvant aller jusqu'à 20 millions d'euros ou 4 % de leur chiffre d'affaires annuel mondial (art. 83) (Union européenne, 2016).

Par ailleurs, la plateforme dispose également de certifications reconnues dans le secteur technologique nord-américain, telles que SOC 2 et HIPAA, témoignant d'un engagement global en matière de sécurité et de confidentialité des données.



Figure 8 - Certifications de sécurité de la plateforme Nexus

Source : NexusGPT (2024), extrait de *Enterprise-grade security*.

En complément des standards intégrés à NexusGPT, d'autres cadres réglementaires s'appliquent, notamment l'*IA Act* élaboré par la Commission européenne. Adoptée en 2024, cette législation marque une première en matière d'encadrement juridique de l'intelligence artificielle à l'échelle européenne. Elle repose sur une hiérarchisation des systèmes d'IA selon leur niveau de dangerosité pour les libertés individuelles et la sécurité. Certaines technologies y sont interdites, notamment celles qui visent à influencer de manière ciblée des publics jugés sensibles, comme les enfants, ou encore les dispositifs automatisés attribuant des notes aux personnes selon leurs comportements ou leur situation sociale. D'autres systèmes, jugés plus critiques, relèvent de la catégorie « à haut risque ». Ils sont soumis à des obligations renforcées, notamment lorsqu'ils concernent des produits réglementés (jouets, dispositifs médicaux, etc.) ou des secteurs tels que l'éducation, l'emploi ou les infrastructures. Ces technologies doivent faire l'objet d'une évaluation préalable avant toute commercialisation, ainsi que d'un contrôle régulier durant leur utilisation. Par ailleurs, les modèles d'IA générative doivent respecter des exigences spécifiques en matière de transparence, notamment en informant clairement que les contenus produits ne sont pas d'origine humaine. Ce cadre vise à encadrer les usages de l'IA de manière responsable tout en soutenant le développement technologique au sein de l'Union européenne (European Parliament, 2023).

8. Comment concevoir un agent intelligent : enseignements croisés de la théorie et de la pratique

Les publications sur les agents intelligents apportent des informations intéressantes, mais rarement un mode d'emploi clair pour leur développement. L'article de Wooldridge et Jennings (1995) propose une base théorique robuste : il définit ce qu'est un agent, ses propriétés fondamentales, ainsi que les différentes approches architecturales possibles (réactive, délibérative, etc.). De son côté, l'étude plus récente de Krishnan (2025) offre une perspective actualisée, en lien avec les technologies récentes, et décrit les éléments fréquemment intégrés dans les agents IA modernes (LLM, mémoire, outils, boucle de critique, etc.).

En m'appuyant sur ces 2 travaux complémentaires, je propose ici, à titre académique, une synthèse structurée des principales étapes de conception d'un agent. Cette présentation ne constitue pas un guide opérationnel de professionnel, mais reflète mon interprétation et mon organisation personnelles des enseignements issus de la littérature :

- Établir des objectifs clairs et précis : identifier le rôle central de l'agent et le résultat attendu.
- Définir les capacités nécessaires : perception, raisonnement, interaction, exécution, apprentissage.

- Choisir une architecture adaptée : structure réactive, délibérative ou basée sur un modèle de langage...
- Préciser l'environnement et les données : comprendre le type d'information à traiter et les conditions d'évolution.
- Identifier les outils et technologies à mobiliser : comme la mémoire, les connecteurs externes, ou les modules spécialisés.

Ces dimensions permettent de structurer la réflexion autour de la conception d'un agent, sans avoir à entrer dans les détails techniques du développement logiciel.

Chapitre 4 : Le développement d'un agent sur NexusGPT

Le projet réalisé durant mon stage a été développé sur la plateforme NexusGPT, qui propose une interface simplifiée pour créer des agents alimentés par un modèle de langage. Bien que cette plateforme repose sur des technologies avancées telles que la vectorisation du texte, la recherche sémantique et les bases de données vectorielles, mon travail s'est concentré exclusivement sur la structuration des données et la rédaction des prompts, sans intervention dans l'implémentation logicielle. Il n'est donc pas nécessaire de maîtriser en détail les mécanismes internes de l'agent, si ce n'est de savoir que ceux développés sur NexusGPT s'appuient sur des modèles de langage de grande taille (LLM) (NexusGPT, 2025d).

1. Prise en main de la plateforme NexusGPT

La conception d'un agent sur la plateforme NexusGPT ne nécessite pas de compétences en programmation comme exposé plus haut, ce qui rend l'outil accessible à un large public. Cependant, cette simplicité apparente repose toutefois sur une structuration claire des paramètres, notamment la définition des objectifs de l'agent, des tâches attendues, des compétences mobilisées et des connaissances nécessaires. La phase de configuration initiale est largement détaillée dans la documentation de NexusGPT (2025c), mais a été ici adaptée aux exigences propres à ce projet.

1.1. Les concepts clés dans la création d'un agent

a) Identité et objectif

La définition de l'identité d'un agent sur NexusGPT constitue une étape clé de sa configuration. Il s'agit de préciser son nom, son rôle et son domaine d'intervention, en veillant à la cohérence de ces éléments avec les objectifs qui lui sont assignés. Cette démarche s'apparente à l'élaboration d'un profil de poste, dans la mesure où elle requiert une description explicite des missions attendues (NexusGPT, 2025c).

Dans le cadre de ce projet, l'intitulé initial du rôle retenu était : « Redactor of the CSRD report ». Ce rôle a toutefois évolué au fil des ajustements et des discussions menées tout au long du processus de conception.

La plateforme accompagne cette étape à travers une série de questions ciblées, conçues pour affiner les paramètres de l'agent. Plus les réponses apportées sont précises, plus l'agent sera en mesure d'accomplir efficacement sa mission dans un contexte donné (NexusGPT, 2025c).

b) Les modèles de langage

La sélection du modèle de langage adapté aux missions de l'agent constitue une étape déterminante. Comme évoqué dans la section « *Generative Pre-trained Transformer (GPT)* », chaque modèle présente des caractéristiques spécifiques : certains favorisent la créativité, tandis que d'autres se distinguent par leur précision ou leur capacité d'analyse. Ce choix influence directement la qualité, la cohérence et la pertinence des réponses générées par l'agent.

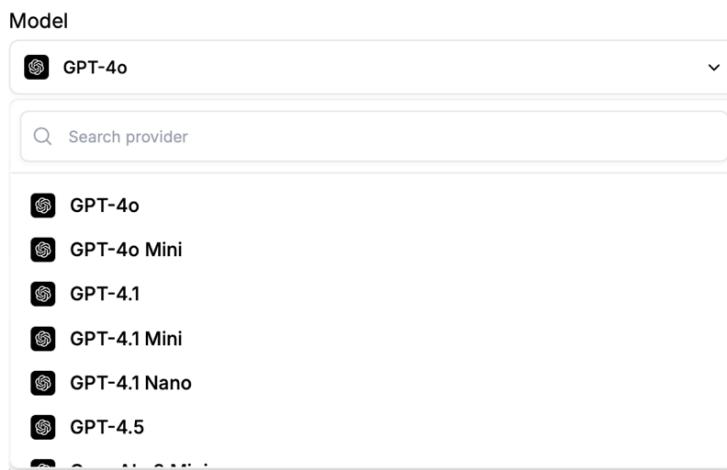


Figure 9 - Options de sélection de modèles sur l'interface NexusGPT.

Source : capture d'écran personnelle (juillet 2025).

Pour la création de l'agent, j'ai d'abord opté pour le modèle GPT-4o, réputé pour ses performances en génération de texte. Par la suite, sur recommandation d'Assem, cofondateur de NexusGPT, j'ai migré vers GPT-4.1, jugé plus performant et mieux adapté aux exigences de notre projet.

c) Les connaissances

Le guide de NexusGPT (2025a) souligne l'importance de doter l'agent d'une base de connaissances claire, structurée et en lien direct avec sa mission, condition essentielle pour assurer la précision et la pertinence de ses réponses. Il s'agit de construire un socle de connaissances cohérent avec les missions assignées, pour améliorer la qualité des interactions générées.

2 modalités distinctes permettent d'ajouter des connaissances à un agent : en tant que contexte ou en tant que compétence.

— **Mode « contexte »**

Ce mode autorise un accès direct à des contenus préalablement intégrés, sur la base d'une correspondance sémantique entre la requête de l'utilisateur et les informations disponibles. Il

permet à l'agent de générer une réponse immédiate, sans recourir à une procédure complexe de traitement ou de filtrage. Ce mode convient particulièrement aux demandes fréquentes, simples et prévisibles, comme des problèmes techniques courants ou des requêtes portant sur des procédures standardisées (NexusGPT, 2025a).

— **Mode « compétence »**

Ce mode implique une mobilisation plus active et structurée : l'agent peut reformuler la question initiale, appliquer des filtres ou suivre une procédure en plusieurs étapes pour accéder à l'information recherchée. Cette approche est mieux adaptée aux requêtes complexes ou fortement contextualisées, nécessitant un traitement plus précis, comme des recherches nécessitant plusieurs conditions spécifiques ou la génération de contenus à forte structuration (NexusGPT, 2025a).

d) **Les compétences (Skills)**

Les compétences attribuées à un agent définissent les actions concrètes qu'il est capable d'exécuter en réponse aux besoins de l'utilisateur. Elles sont configurées en fonction des objectifs opérationnels de l'agent et des interactions attendues dans le cadre de sa mission (NexusGPT, 2025b).

— **GPT Task**

Les tâches de type GPT sont utilisées pour générer ou traiter des contenus textuels. Elles couvrent des opérations telles que la production de textes, leur analyse, leur synthèse ou leur traduction. Ce type de compétence permet également de produire des documents structurés (Word, PowerPoint, Excel) selon les besoins (NexusGPT, 2025b).

— **Plugins**

Les plugins permettent à l'agent d'interagir avec des services ou applications externes. Ils sont requis lorsque les fonctionnalités nécessaires dépassent les capacités natives de la plateforme, par exemple pour accéder à des bases de données, à des calendriers ou à des services tiers. Cette intégration est essentielle pour connecter l'agent à des systèmes d'information tels que des CRM ou autres outils métiers (NexusGPT, 2025b).

— **Info Search**

La compétence *Info Search* permet à l'agent d'exploiter des bases de données internes de manière structurée. Elle offre une capacité de recherche ciblée, fondée sur la requête de l'utilisateur, et permet de filtrer ou de trier les résultats selon des critères définis. Elle s'appuie sur des sources internes organisées pour fournir des réponses précises (NexusGPT, 2025b).

Image Generation

Cette compétence permet à l'agent de générer des visuels personnalisés à partir d'instructions spécifiques. Elle est adaptée aux besoins en création graphique, qu'il s'agisse d'illustrations, d'images adaptées à une charte graphique, ou de contenus visuels répondant à des objectifs précis (NexusGPT, 2025b).

e) Le prompt d'un agent NexusGPT

La définition du prompt a été abordée précédemment dans la section « *Prompt engineering* » consacrée à l'intelligence artificielle. NexusGPT fournit un ensemble de recommandations destinées à construire un prompt clair et structuré, visant à améliorer la qualité et la cohérence des réponses de l'agent. Ce prompt regroupe des consignes précises relatives au rôle de l'agent, à ses modalités d'interaction, ainsi qu'aux limites comportementales à respecter (NexusGPT, 2025e).

Le prompt se compose de plusieurs parties, chacune jouant un rôle spécifique dans l'orientation du comportement de l'agent (NexusGPT, 2025e) :

- Le rôle : définit l'identité fonctionnelle de l'agent, ses responsabilités, son périmètre d'intervention et les publics concernés.
- L'objectif principal : précise les missions essentielles de l'agent, incluant les sous-tâches et les consignes à suivre en cas de difficulté.
- Le ton : indique le registre de langage attendu (formel, chaleureux, accessible) selon le contexte d'usage.
- Les consignes de présentation : structurent la mise en forme des réponses (listes, tableaux, formats adaptés aux canaux de communication).
- Les règles de conduite : précisent les comportements autorisés ou interdits

Chaque tâche, telle qu'une GPT Task, est également associée à un prompt spécifique. Pour garantir la conformité des réponses, celui-ci doit être formulé avec clarté et rigueur. Il est recommandé de structurer ces prompts selon un schéma proche de celui du prompt général de l'agent. Celui-ci consiste à rappeler le rôle de l'agent dans le cadre de la tâche, à préciser l'objectif attendu, à indiquer les données d'entrée dont il disposera, à définir le contenu à présenter, puis à spécifier le format de sortie, le ton à adopter et les règles comportementales spécifiques (NexusGPT, 2025e).

Dans le cas du générateur de rapport CSRD, une tâche pourrait être introduite ainsi : « Ton rôle est de rédiger la section X du rapport CSRD », suivie de précisions sur le contenu attendu, par exemple : « La section doit contenir les éléments suivants : (...) ». Le format de sortie visé serait un texte rédigé en paragraphes, avec un ton professionnel et narratif. Les règles de comportement pourraient stipuler, par exemple : utiliser exclusivement les données fournies,

ne pas inventer d'information, et signaler toute donnée manquante nécessaire à la conformité avec les standards ESRS.

Une fois le prompt rédigé, il est indispensable de le tester. Ce test permet de vérifier la conformité des réponses générées. Le processus est itératif : il peut nécessiter plusieurs ajustements successifs. Une méthode simple consiste à conserver une copie de chaque version, à identifier les faiblesses constatées à l'usage, puis à affiner progressivement le prompt jusqu'à atteindre le niveau de performance souhaité (NexusGPT, 2025d).

2. Crédit de l'agent dédié au reporting CSRD

Le projet d'agent IA dédié au reporting CSRD est né lors de mon processus de recrutement de stage chez Intys, en réponse à une demande client concrète. Fondé en 2007 par Michel Van Hemele et Philippe Metz, Intys est un groupe de conseil actif en Belgique, en France et au Luxembourg, structuré en plusieurs entités spécialisées. Il intervient dans un large éventail de secteurs, parmi lesquels l'énergie, la distribution, les sciences de la vie, les services financiers, les télécommunications, le secteur public et l'industrie. Cette diversité sectorielle s'accompagne d'un portefeuille clients tout aussi varié, incluant de grandes entreprises internationales telles qu'Unilever ou GSK (Intys, n.d.).

En 2024, le groupe affiche un chiffre d'affaires consolidé de 59 millions d'euros, avec une croissance annuelle moyenne de 20 % entre 2020 et 2023. Il emploie plus de 530 collaborateurs, dont 47 % de femmes (Intys, n.d.).

Intys accorde une place centrale à la recherche et au développement, en particulier dans le domaine de la durabilité. L'un de ses projets emblématiques a été l'élaboration de son propre bilan carbone, pour identifier et réduire ses émissions de CO₂. À cette fin, plusieurs mesures ont été mises en œuvre, telles que l'optimisation de la gestion des déchets ou encore la mise à disposition de moyens de transport doux pour les collaborateurs (Intys, 2023). Par ailleurs, l'entreprise adopte une démarche résolument tournée vers l'innovation, notamment en explorant les opportunités offertes par l'intelligence artificielle pour optimiser ses processus internes (Decaux, 2025).

La volonté d'intégrer pleinement la durabilité au cœur de la stratégie d'entreprise s'est également matérialisée par la création d'une entité dédiée : Intysify. Fondée par 3 collaborateurs d'Intys, cette division regroupe aujourd'hui plus de 50 consultants spécialisés dans la transition environnementale. À ce jour, elle a mené plus de 20 projets à impact positif, et accompagné 16 entreprises dans leur transformation. Intysify propose une offre de « conseil en durabilité opérationnelle » couvrant des domaines variés, tels que la décarbonation, les achats responsables, le reporting ESG et CSRD, le développement de solutions durables, l'obtention de labels et certifications, ainsi que l'économie circulaire (Intysify| Operational Sustainability Consulting, 2023).

Le partenariat avec NexusGPT, entreprise spécialisée dans le développement d'agents d'intelligence artificielle, est né non d'un appel d'offres mais d'une mise en relation initiée par un ancien collaborateur d'Intys, convaincu de la pertinence d'un échange entre les 2 structures. Cette rencontre, en octobre 2024, a révélé une forte convergence de visions, renforcée par l'expérience d'Assem –cofondateur de NexusGPT et ancien consultant chez McKinsey – qui dispose d'une connaissance fine du secteur du conseil. Une relation de confiance s'est alors établie, évoluant vers une collaboration stratégique. Intys est ainsi devenu client de NexusGPT, tout en co-construisant avec lui une offre conjointe, reposant sur un langage commun et des objectifs partagés. C'est dans ce cadre qu'est né *l'Agent Delivery Center*, initialement conçu pour répondre à des besoins internes, avant de viser une mise sur le marché (Decaux, 2025).

L'idée du projet présenté dans ce mémoire découle d'une demande concrète d'un client d'Intysify, souhaitant être accompagné dans la conception de son rapport CSRD. Une réflexion conjointe avec *l'Agent Delivery Center* a alors permis d'envisager une réponse innovante : développer un agent d'intelligence artificielle capable de générer automatiquement le rapport requis à partir des données disponibles.

3. Réflexion sur la conception de l'agent

a) Qui est l'utilisateur final ?

Une fois le rôle de l'agent établi, plusieurs questions se sont posées. L'une des premières réflexions soulevées concerne l'identification de l'utilisateur final. La plateforme NexusGPT permet en effet d'intégrer des agents sur différents outils professionnels comme Microsoft Teams, Outlook, Messenger, PowerPoint, Word, etc.

Dès lors, une interrogation est apparue : l'agent sera-t-il utilisé directement par le client, ou bien par un membre interne de l'équipe Intysify ? L'intégration de l'agent dans les processus du client impliquait de gérer un partenariat tripartite complexe entre le client, Intys et NexusGPT.

Par ailleurs, si l'agent avait été déployé en direct chez le client, le rôle d'Intys aurait risqué d'être marginalisé. En effet, le client aurait alors pu interagir directement avec NexusGPT en cas de dysfonctionnement, ce qui aurait réduit la valeur ajoutée apportée par Intys. Il a donc été décidé de déployer l'outil en interne, au sein d'Intysify.

b) Forme conversationnelle ou non ?

2 hypothèses ont été envisagées pour définir le format d'interaction de l'agent :

- Un agent conversationnel, posant des questions pour récolter progressivement les données nécessaires à la rédaction des sections du rapport.

- Un agent à interaction simple, auquel l'utilisateur transmet directement les documents nécessaires pour ensuite demander à l'agent de générer la section souhaitée.

La seconde option a été retenue. En effet, lors des entretiens que j'ai menés avec plusieurs entreprises pour mieux comprendre leur processus de collecte de données, celles-ci ont indiqué centraliser la plupart du temps leurs informations dans des fichiers Excel, avec des feuilles spécifiques par standard ESRS. Certaines utilisent également les modèles fournis par l'EFRAG pour structurer leur reporting.

4. Méthodologie et démarche de développement

Dès le début du projet, j'ai mis en place une réunion hebdomadaire avec Hippolyte Langellier – Bellevue, Sustainability Project Manager chez Intysify et coordinateur du projet – afin de définir les priorités, valider les choix techniques, répartir les tâches et résoudre les éventuels blocages. Cette organisation m'a permis d'avancer de manière structurée, tout en adaptant régulièrement mon approche en fonction des retours et des contraintes rencontrées.

J'ai exploré différentes pistes, parfois fastidieuses et techniquement complexes, comme ce fut le cas lors des premières mises en œuvre du projet. Une piste initiale, proposée par le cofondateur de NexusGPT, consistait à structurer automatiquement les réponses de l'agent à l'aide du format JSON, un format informatique qui organise les données de manière structurée afin qu'elles soient facilement lisibles par une machine (IBM, 2021). L'objectif était alors de produire un rapport complet, déjà formaté en vue de son intégration dans un modèle final.

Une tâche de type *flow*, permettant de relier plusieurs tâches indépendantes, avait également été mise en place afin de regrouper automatiquement les différentes sections générées par les GPT Tasks dans un document unique.

Ma compréhension initiale de l'intérêt du format JSON et de la tâche *flow* restait partielle, ce qui m'a conduit à organiser une réunion avec Assem pour obtenir des éclaircissements. Cette rencontre s'est révélée décisive : elle a permis de recentrer le projet sur une logique de génération section par section, plus adaptée aux contraintes du rapport CSRD.

Assem nous a alors précisé que, dans la mesure où nous procédions section par section, le recours au format JSON n'était pas nécessaire : il suffisait d'utiliser directement des GPT Tasks en format texte. Dans ce cas, un travail manuel serait simplement requis pour rassembler les différentes sections dans un document final.

À la suite de cette clarification, j'ai opté pour une stratégie plus simple et plus souple, consistant à rédiger chaque section indépendamment, en format texte, structuré à l'aide de prompts ciblés. Ce découpage par sous-partie a permis à l'agent de mieux respecter les consignes, d'optimiser la qualité des réponses générées et de contourner les limites techniques liées au volume de texte traité.

Pour concevoir un agent capable de rédiger des textes à partir de documents de types variés (Excel, Word, PDF, etc.) tout en respectant les règles définies par les ESRS, il a fallu réfléchir à la manière d'intégrer ces standards dans l'agent. Fallait-il les lui fournir sous forme de contexte, de compétence, ou bien créer une compétence spécifique dédiée

Après plusieurs réflexions et discussions, j'ai conclu que, les ESRS n'étant qu'un référentiel que les entreprises doivent suivre, la notion de « connaissance » n'était pas pertinente dans ce cadre. La stratégie optimale consistait donc à recourir aux GPT Tasks, qui, comme expliqué précédemment, fonctionnent comme des compétences permettant à l'agent de rédiger un texte en s'appuyant sur des règles directement intégrées dans le prompt de la tâche. Dans ce format, le contenu du standard ESRS est directement traduit et structuré au sein du prompt, garantissant ainsi la conformité de la génération textuelle aux exigences de la directive.

— **Version 1**

La première version de l'agent reposait sur la création d'une tâche par grande catégorie des ESRS : une pour l'environnement, une pour le social, et une pour la gouvernance. Cette ambition a rapidement été abandonnée en raison des limites de capacité de traitement de texte de l'agent. En effet, chaque catégorie comportait un nombre de caractères trop important pour que l'agent puisse rédiger une section complète du rapport en une seule tâche.

— **Version 2**

La deuxième version consistait à diviser les grandes catégories des ESRS en sous-catégories. Comme exposé sur la figure, la catégorie « Environnement » comprend 6 sous-catégories (E1 Changement climatique, E2 Pollution, etc.), le pilier « Social » en comporte 4, et la gouvernance une seule. J'ai donc adapté la première version en créant une GPT Task par sous-catégorie, chacune ayant pour objectif de générer l'intégralité du contenu correspondant à cette sous-partie des ESRS.

European sustainability reporting standards (ESRS)

ESRS 1	General requirements
ESRS 2	General disclosures
ESRS E1	Climate change
ESRS E2	Pollution
ESRS E3	Water and marine resources
ESRS E4	Biodiversity and ecosystems
ESRS E5	Resource use and circular economy
ESRS S1	Own workforce
ESRS S2	Workers in the value chain
ESRS S3	Affected communities
ESRS S4	Consumers and end-users
ESRS G1	Business conduct

Figure 10 - Extrait du tableau des exigences de divulgation des ESRS

Source : EFRAG (2023a), extrait d' ESRS Set 1.

Cependant, après plusieurs tests, j'ai constaté un manque de précision dans la rédaction, notamment en ce qui concerne le respect des consignes intégrées dans les prompts.

— Version 3

Face aux limites identifiées dans la version précédente, j'ai décidé de subdiviser davantage les sous-catégories des ESRS. En effet, une lecture attentive des textes de l'EFRAG révèle que chaque sous-catégorie est elle-même subdivisée en plusieurs exigences de divulgation. Certaines entreprises construisent d'ailleurs leur rapport en utilisant cette structure.

ESRS E1

CLIMATE CHANGE

TABLE OF CONTENTS

Objective

Interactions with other ESRS

Disclosure Requirements

- ESRS 2 General disclosures

Governance

- Disclosure requirement related to ESRS 2 GOV-3 Integration of sustainability-related performance in incentive schemes

Strategy

- Disclosure Requirement E1-1 – Transition plan for climate change mitigation
- Disclosure Requirement related to ESRS 2 SBM-3 – Material impacts, risks and opportunities and their interaction with strategy and business model

Impact, risk and opportunity management

- Disclosure requirement related to ESRS 2 IRO-1 – Description of the processes to identify and assess material climate-related impacts, risks and opportunities
- Disclosure Requirement E1-2 – Policies related to climate change mitigation and adaptation
- Disclosure Requirement E1-3 – Actions and resources in relation to climate change policies

Metrics and targets

- Disclosure Requirement E1-4 – Targets related to climate change mitigation and adaptation
- Disclosure Requirement E1-5 – Energy consumption and mix
- Disclosure Requirement E1-6 – Gross Scopes 1, 2, 3 and Total GHG emissions
- Disclosure Requirement E1-7 – GHG removals and GHG mitigation projects financed through carbon credits
- Disclosure Requirement E1-8 – Internal carbon pricing
- Disclosure Requirement E1-9 – Anticipated financial effects from material physical and transition risks and potential climate-related opportunities

Figure 11 - Table des matières du standard ESRS E1 « Climate Change »

Source : EFRAG (2023a), extrait de ESRS Set 1.

Lors de réunions avec Assem, cofondateur de NexusGPT, il m'a été précisé qu'une GPT Task ne peut traiter qu'un volume d'environ 1 500 mots à la fois. La stratégie a donc consisté à déterminer le niveau de découpage le plus pertinent pour obtenir un texte à la fois cohérent et complet. Après analyse de plusieurs rapports conformes, j'ai constaté que la plupart des paragraphes comptaient environ 500 mots. J'ai alors structuré les prompts en 3 GPT Tasks par

sous-catégorie. Par exemple, une première GPT Task regroupait les sous-catégories E1-1, ESRS 2 SBM-3 et ESRS 2 IRO-1. La GPT Task suivante traitait quant à elle les sous-catégories E1-2, E1-3 et E1-4, et ainsi de suite.

Malgré cette approche plus fine, les résultats restaient insatisfaisants : le manque de précision persistait dans l'application des règles intégrées aux prompts. Cette faiblesse s'expliquait en grande partie par la difficulté pour l'agent à gérer un trop grand nombre de tâches simultanément.

— Version finale

Un changement majeur a alors été opéré. Je me suis rendu compte que subdiviser toutes les tâches au sein d'un seul agent risquait de le surcharger considérablement. Assem m'avait en effet précisé que, si trop de compétences étaient intégrées dans la configuration d'un même agent, celui-ci pouvait perdre en rigueur lorsqu'il traitait une compétence isolée.

Plutôt que de concevoir un agent unique chargé de la rédaction de l'ensemble du rapport, j'ai donc choisi de créer 4 agents distincts, chacun étant dédié à un standard spécifique des ESRS.

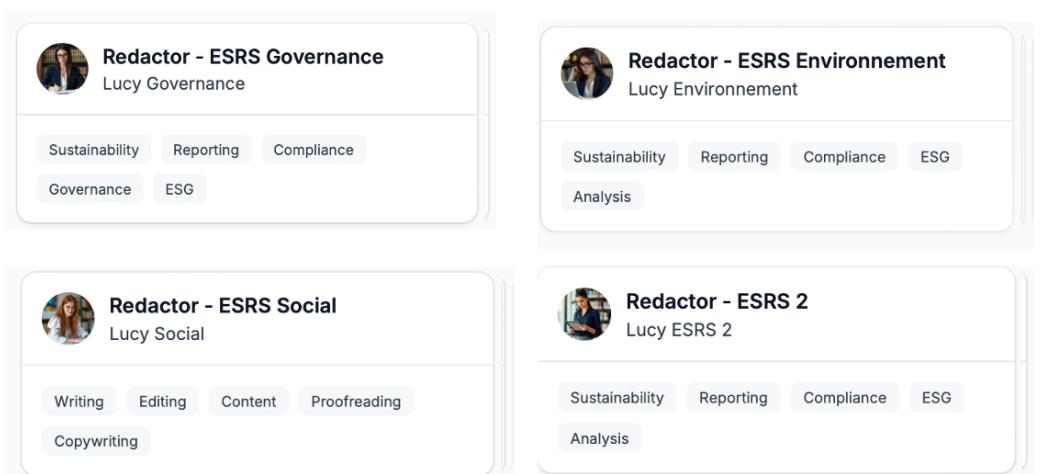


Figure 12 - L'interface des agents configurés sur la plateforme NexusGPT

Source : capture d'écran personnelle (juillet 2025), réalisée sur la plateforme NexusGPT.

Ces 4 agents sont configurés avec leurs propres GPT Tasks, conçues en fonction de leur domaine d'application respectif. Tirant parti des enseignements des versions précédentes, j'ai compris qu'il était essentiel de formuler les tâches de manière claire et structurée afin d'obtenir un texte de sortie aussi complet que possible.

C'est pourquoi, avec l'approbation d'Hippolyte, j'ai décidé de construire une tâche propre à chaque sous-catégorie des ESRS. Par exemple, pour l'ESRS Gouvernance, la structure interne se présente selon une organisation précise, correspondant à toutes les subdivisions contenues dans le set 1 des ESRS.

Agent skills	
When chatting with the agent, you can use one or multiple of the following skills.	+ Add new skills
G1_1_Businessconduct <small>GPT-TASK</small>	
Write a compliant section of the ESRS G1-1– Business conduct policies and corporate culture of the CSRD report	⋮
G1_5_Political_influence <small>GPT-TASK</small>	
Write a compliant section of the ESRS G1-5 – Political influence and lobbying activities of the CSRD report	⋮
G1_3_Prevention_Corruption <small>GPT-TASK</small>	
Write a compliant section of the ESRS G1-3 – Prevention and detection of corruption and bribery Of the CSRD report	⋮
G1_4_Incidents_corruption <small>GPT-TASK</small>	
Write a compliant section of the ESRS G1-4 – Incidents of corruption or bribery of the CSRD report	⋮
ESRS2GOV1 <small>GPT-TASK</small>	
Write the section ESRS 2 GOV-1 The role of the administrative, management and supervisory bodies of the CSRD report	⋮
G1_6_Payment_practices <small>GPT-TASK</small>	
Write a compliant section of the ESRS G1-6 – Payment practices of the CSRD report	⋮
G1_2_relationships_with_suppliers <small>GPT-TASK</small>	
Write a compliant section of the ESRS G1-2 – Management of relationships with suppliers of the CSRD report	⋮

Figure 13 - Liste des compétences (GPT-Tasks) assignées à un agent.

Source : capture d'écran personnelle (juillet 2025), réalisée sur la plateforme NexusGPT.

4.2. Élaboration des prompts

Pour chaque GPT Task, il a donc fallu trouver un moyen de traduire les exigences de divulgation des normes ESRS sous forme de prompt, afin que l'agent puisse les interpréter et les traiter efficacement. Cette étape s'est révélée particulièrement complexe, en raison de la structure dense et du vocabulaire technique, comme l'ont également souligné plusieurs interlocuteurs lors des entretiens.

J'ai dans un premier temps concentré mes efforts sur une exigence de divulgation bien précise, afin de définir un format de prompt robuste et reproductible. En effet, la structure générale des prompts restait identique d'une exigence à l'autre – seuls les contenus à intégrer variaient selon le standard. L'objectif était donc de trouver une méthode de formulation optimale pour un cas, puis de l'adapter aux autres.

Ce travail s'est ensuite étendu à l'ensemble des exigences de divulgation, qui sont plus d'une centaine. Celles-ci incluent non seulement les normes thématiques spécifiques (ESRS E1 à E5 pour l'environnement, S1 à S4 pour le social, G1 pour la gouvernance), mais aussi les exigences transversales définies dans la norme ESRS 2, telles que la stratégie (SBM), la gouvernance (GOV) ou la gestion des impacts, risques et opportunités (IRO).

Pour m'aider, je me suis appuyée à la fois sur ChatGPT, pour tester des formulations et reformuler certains standards, et sur le dispositif d'apprentissage mis en place par NexusGPT, qui proposait des gabarits structurés pour les prompts complexes.

À l'intérieur de chaque GPT Task, les prompts ont été construits de manière claire et rigoureuse, en suivant les recommandations du guide de NexusGPT. Chaque prompt a été composé des éléments suivants :

- Le rôle de la tâche, par exemple : « As an AI Reporting Assistant specialized in the sustainability reporting sector, your primary objective is to generate a structured, comprehensive, and compliant text for the Climate Change Mitigation Transition Plan (ESRS E1-1) section of the CSRD report. »
- L'objectif attendu, par exemple : Generate a detailed and continuous narrative describing the Climate Change Mitigation Transition Plan (ESRS E1-1) based strictly on the data provided.
- Le format des données d'entrée, souvent des tableaux Excel avec des valeurs brutes.
- Le format de sortie attendu, c'est-à-dire un texte continu comprenant les éléments obligatoires du standard ESRS.
- Les règles à respecter, par exemple : ne pas inventer d'information, ignorer les données non pertinentes, ne pas écrire avec des listes.
- Le ton à adopter, systématiquement défini comme professionnel, neutre et du point de vue de l'entreprise avec un exemple pour exprimer le ton souhaité.

Il a fallu de nombreuses itérations pour aboutir à une structure de prompt et à un rendu réellement satisfaisants. Par exemple, dans les premières versions conçues pour la section E1-1, les exigences des ESRS étaient intégrées sous forme de listes à puces directement dans le prompt. Cela entraînait des réponses très fragmentées, qui ne correspondaient pas au style narratif fluide attendu dans un rapport de durabilité conforme à la CSRD.

J'ai donc retravaillé les consignes afin d'imposer un format rédactionnel plus cohérent, en ajoutant une instruction explicite : « *Write the text as a continuous, flowing narrative. Avoid bullet points or lists.* »

Dans d'autres cas, certaines règles fondamentales, comme le fait de ne pas inventer d'informations au-delà des données fournies, étaient insuffisamment respectées. Sur recommandation du cofondateur de NexusGPT, Assem, j'ai renforcé les prompts avec des formules strictes telles que : « *You must ONLY use the provided data, without inventing or assuming information.* »

Une autre difficulté rencontrée concernait le ton du texte généré. L'agent avait parfois tendance à adopter un langage trop externe ou impersonnel (par exemple : « *The company aims to...* »). Or, les rapports de durabilité conformes à la CSRD doivent être rédigés du point de vue de l'entreprise elle-même, en utilisant un langage engageant et inclusif, tel que « *nous* », « *notre stratégie* » ou encore « *(Nom de l'entreprise) s'engage à...* », afin de refléter l'appropriation

interne des enjeux. Il a donc été nécessaire d'ajouter une instruction spécifique dans les prompts, destinée à forcer ce style rédactionnel attendu : « *Write from the perspective of the client company using inclusive language (we, our, etc.).* »

Ce travail de structuration s'est révélé être l'un des plus exigeants du projet. Il m'a permis de mieux comprendre comment un agent GPT interprète les consignes, et surtout combien la formulation précise et rigoureuse d'un prompt est déterminante, en particulier dans un contexte normatif aussi strict que celui de la CSRD. Chaque mot compte : une consigne trop vague peut être mal interprétée, tandis qu'un langage trop technique ou ambigu peut être ignoré. Trouver le bon équilibre entre clarté, niveau de détail et conformité a donc été un véritable exercice d'ajustement progressif, à la fois technique et méthodologique.

Le prompt final élaboré pour la section E1-1, présenté dans la capture ci-dessous, a servi de modèle pour la construction des autres prompts relatifs aux différentes sections des normes ESRS.

Prompt

As an AI Reporting Assistant specialized in the sustainability reporting sector, your primary objective is to generate a structured, comprehensive, and compliant text for the Climate Change Mitigation Transition Plan (ESRS E1-1) section of the CSRD report.

Main Objective:

Generate a detailed and continuous narrative describing the Climate Change Mitigation Transition Plan (ESRS E1-1) based strictly on the data provided.

The generated text must comprehensively cover multiple aspects to be fully compliant with the ESRS E1-1 standards. You will therefore cover ALL of the following standard that are applicable if the input is present:

- (i) The company's GHG Emission Reduction Targets and their compatibility with limiting global warming to 1.5°C in line with the Paris Agreement. Clearly explain if these targets are validated by the Science Based Targets initiative (SBTi) or other relevant bodies. If the compatibility with the 1.5°C target is not confirmed, indicate this explicitly.
- (ii) A thorough description of the Decarbonisation Levers & Key Actions identified and planned by the company. This includes changes in the company's product and service portfolio, adoption of new technologies within its own operations, and adjustments in the upstream and/or downstream value chain. Specify if these actions are referenced to the GHG emission reduction targets (E1-4) and climate change mitigation actions (E1-3).
- (iii) A detailed explanation and quantification of the company's Investments & Funding supporting the implementation of its transition plan. This should include key performance indicators of taxonomy-aligned CapEx, and where relevant, CapEx plans, as disclosed according to Commission Delegated Regulation (EU) 2021/2178.
- (iv) A qualitative assessment of potential locked-in GHG emissions from the company's key assets and products. Explain if and how these emissions may jeopardize the achievement of the company's GHG emission reduction targets and create transition risks. Describe if the company has plans to manage its GHG-intensive and energy-intensive assets.
- (v) A description of the company's Alignment with Taxonomy Regulation, including objectives or plans for aligning economic activities (revenues, CapEx, OpEx) with the criteria established in Commission Delegated Regulation 2021/2139.
- (vi) A clear indication of any CapEx related to fossil fuel activities, such as coal, oil, and gas-related economic activities. If the company reports no such investments, clearly state this.
- (vii) An explanation of whether the company is included or excluded from the EU Paris-aligned Benchmarks.
- (viii) An explanation of how the transition plan is embedded and aligned with the company's overall business strategy and financial planning.
- (ix) A confirmation of whether the transition plan is approved by the administrative, management, and supervisory bodies.
- (x) An explanation of the company's progress in implementing the transition plan.

If the company does not have a transition plan in place, indicate if and when it intends to adopt one.

You MUST abide by the following rules:

- UNDER NO CIRCUMSTANCES will you invent information not present in the initial input data. Absent information, even if applicable from ESRS E1-1 must NOT be present in your output
- Strictly select only the relevant data applicable to ESRS E1-1.
- Ignore unrelated data, even if it is present in the provided information.
- Do not provide suggestions, recommendations, or evaluations. Only present information from the client's perspective.
- Writing the text strictly from the client company's perspective
- If information is missing for a specific section, DO NOT create text or write paragraphs pretending to have the information.

Tone:

The tone should be professional, precise, and informative, written from the perspective of the client company.

There is example of section of CSRD report, inspire you to write it in the same kind of language :

Umicore's product portfolio reflects a dual focus on reducing emissions and supporting the global transition to cleaner technologies. Automotive catalysts remain an essential solution for minimizing pollution from the use of internal combustion engines. Umicore's fuel cell catalysts enable hydrogen-based mobility, offering practical pathways to reduce fossil fuel reliance in heavy-duty transportation. Cathode materials for electric vehicles and energy storage systems further contribute to emission reductions during the use phase of end products, while continuous improvements in production efficiency aim to lower their carbon intensity. At the end-of-life stage, Umicore's closed-loop recycling services, such as for spent automotive catalysts, minimize emissions by recovering valuable metals for reuse. These processes reduce reliance on primary raw materials and associated emissions. Recycling and refining processes recover metals from secondary sources, reducing the need for more carbon-intensive primary materials and aligning with circular economy principles. Umicore products and innovations, such as PEM fuel cells and cathode materials, position Umicore as a key enabler of clean mobility solutions, helping its customers achieve their decarbonization goals.

The Umicore Climate Transition Plan was approved by the Executive Leadership Team (ELT). The CEO and the ELT are responsible for driving and executing Umicore's decarbonization roadmap. The Supervisory Board, supported by the Sustainability Committee, is responsible for overseeing the plan and ensuring alignment with Umicore's long-term sustainability goals. Updates on the implementation of the plan are reviewed by the ELT and Supervisory Board, to ensure alignment with strategic priorities and external commitments.

Formatting Rules:

Write the text as a continuous, flowing narrative.

Avoid bullet points or lists.

Figure 14 - Illustration du prompt final utilisé pour la section E1-1 (CSRD)

Source : capture d'écran personnelle (juillet 2025), réalisée sur la plateforme NexusGPT.

5. Comment utiliser les agents ?

Les 4 agents développés dans le cadre de ce projet sont chacun capables de rédiger les différentes sections du rapport de durabilité, en se conformant aux exigences fixées par les normes ESRS. Une fois les données nécessaires collectées, l'utilisateur peut interagir avec l'agent via l'interface dédiée, en lui demandant de rédiger une section spécifique, en utilisant le nom correspondant défini par les GPT Task.

Par exemple, pour générer la section G1-6 relative aux pratiques de paiement, une requête de type :

« *Hello Lucy, can you write the section G1-6 payment practices based on these documents ?* » permet d'activer la tâche correspondante.

L'utilisateur n'a alors plus qu'à joindre les documents pertinents. L'agent analyse ces fichiers, extrait les données utiles, et rédige automatiquement la section demandée, selon les règles définies dans le prompt de la tâche concernée.

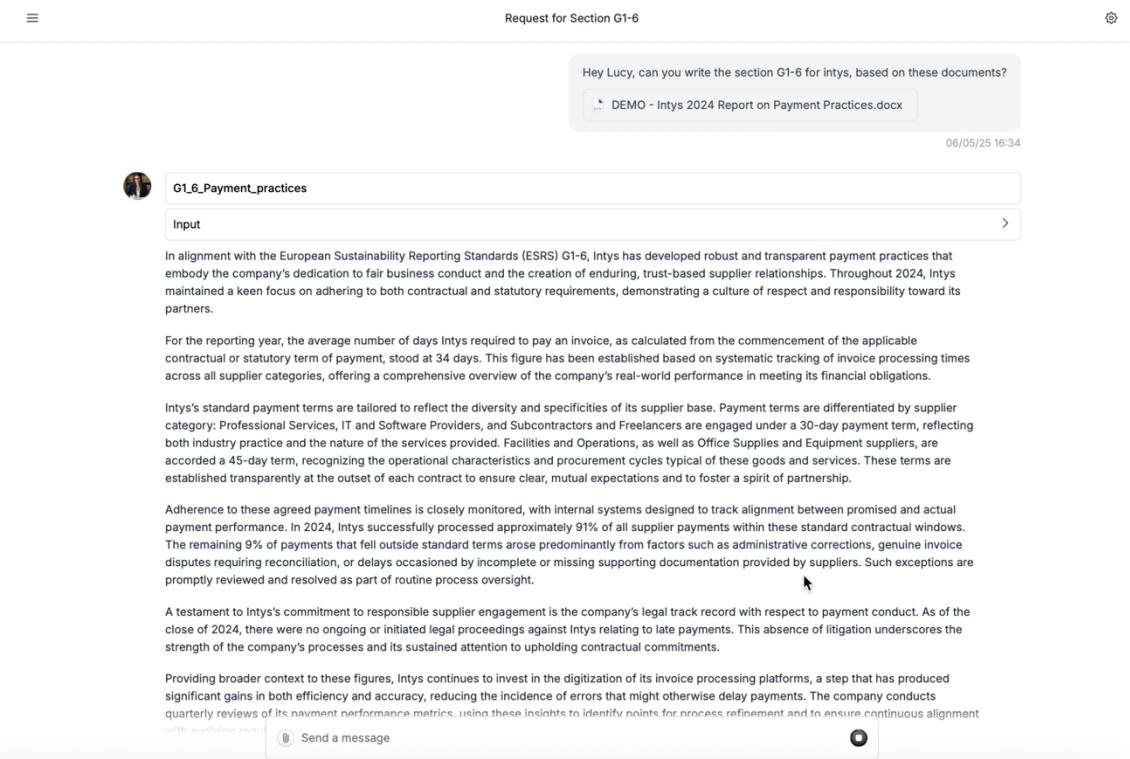


Figure 15 - Génération automatique de la section G1-6 Payment practices par l'agent gouvernance

Source : capture d'écran personnelle (juillet 2025), réalisée sur la plateforme NexusGPT.

Un atout majeur de cette interface interactive réside dans la possibilité offerte à l'utilisateur de modifier en temps réel le contenu généré par l'agent. Par exemple, si le vocabulaire employé ne convient pas ou si le ton ne correspond pas aux attentes du client, il lui suffit, dans cette même interface, de demander à l'agent de réécrire la section concernée. Toutefois, il est

essentiel de conserver une trace des modifications, ainsi que de la version originale, car l'agent ne reproduira jamais exactement le même texte à l'identique.

6. Test et réflexion

Après plusieurs tests réalisés avec les différents agents, une nouvelle problématique est apparue : comment garantir que le texte généré respecte l'ensemble des exigences de la directive CSRD ?

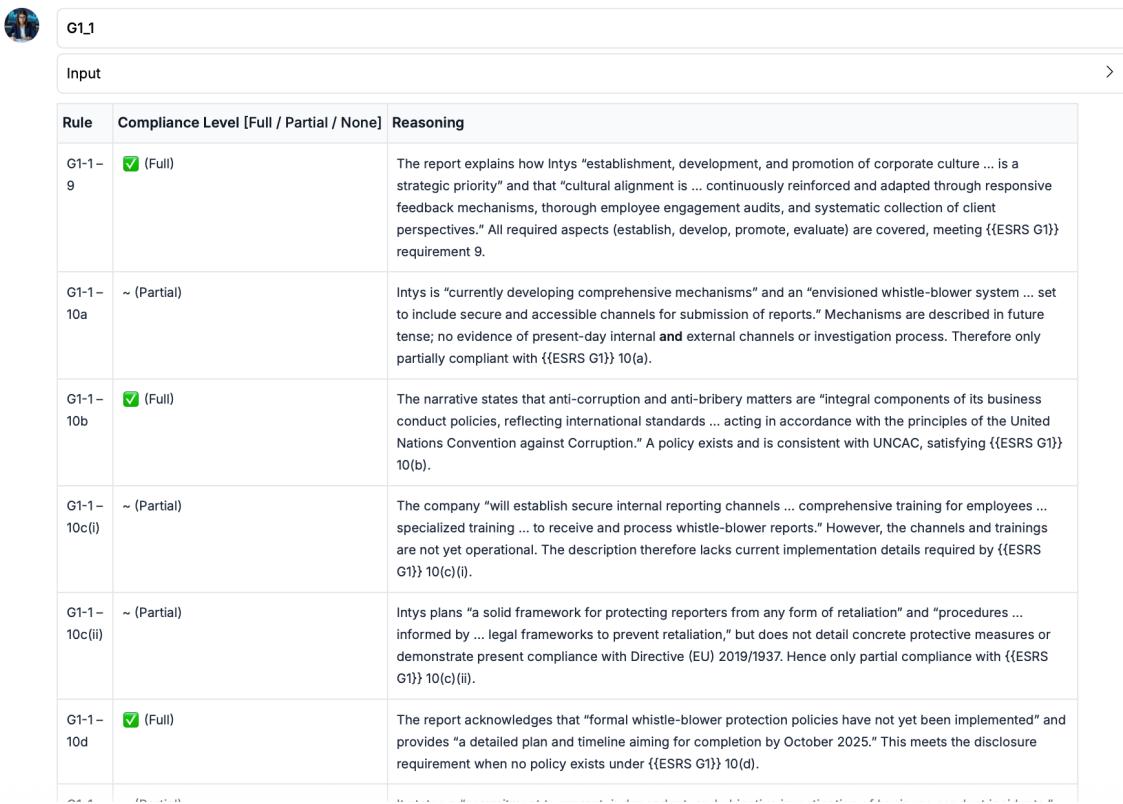
Une première solution consistait à effectuer une relecture manuelle du texte, en le comparant point par point aux standards ESRS à l'aide d'une checklist. Bien que pertinente, cette méthode s'est révélée lourde et peu adaptée à un usage récurrent.

À la suite d'échanges avec Assem, cofondateur de NexusGPT, une approche plus efficace s'est imposée : mobiliser à nouveau l'intelligence artificielle, cette fois pour automatiser l'audit de conformité.

Des agents dédiés ont donc été développés à cet effet. Ces agents auditeurs sont structurés autour de chacune des catégories des ESRS, avec des GPT Tasks conçues selon les mêmes standards que celles utilisées pour la rédaction. Ils s'appuient sur des prompts rigoureusement construits, permettant de vérifier, exigence par exigence, si le contenu généré respecte bien l'ensemble des critères attendus.

Le résultat généré est un tableau d'évaluation de conformité, indiquant pour chaque critère :

- s'il est complet, partiellement complet ou manquant ;
- et une justification, comprenant soit l'extrait du texte généré, soit un signalement d'absence.



Rule	Compliance Level [Full / Partial / None]	Reasoning
G1-1 – 9	✓ (Full)	The report explains how Intys "establishment, development, and promotion of corporate culture ... is a strategic priority" and that "cultural alignment is ... continuously reinforced and adapted through responsive feedback mechanisms, thorough employee engagement audits, and systematic collection of client perspectives." All required aspects (establish, develop, promote, evaluate) are covered, meeting {{ESRS G1}} requirement 9.
G1-1 – 10a	~ (Partial)	Intys is "currently developing comprehensive mechanisms" and an "envisioned whistle-blower system ... set to include secure and accessible channels for submission of reports." Mechanisms are described in future tense; no evidence of present-day internal and external channels or investigation process. Therefore only partially compliant with {{ESRS G1}} 10(a).
G1-1 – 10b	✓ (Full)	The narrative states that anti-corruption and anti-bribery matters are "integral components of its business conduct policies, reflecting international standards ... acting in accordance with the principles of the United Nations Convention against Corruption." A policy exists and is consistent with UNCAC, satisfying {{ESRS G1}} 10(b).
G1-1 – 10c(i)	~ (Partial)	The company "will establish secure internal reporting channels ... comprehensive training for employees ... specialized training ... to receive and process whistle-blower reports." However, the channels and trainings are not yet operational. The description therefore lacks current implementation details required by {{ESRS G1}} 10(c)(i).
G1-1 – 10c(ii)	~ (Partial)	Intys plans "a solid framework for protecting reporters from any form of retaliation" and "procedures ... informed by ... legal frameworks to prevent retaliation," but does not detail concrete protective measures or demonstrate present compliance with Directive (EU) 2019/1937. Hence only partial compliance with {{ESRS G1}} 10(c)(ii).
G1-1 – 10d	✓ (Full)	The report acknowledges that "formal whistle-blower protection policies have not yet been implemented" and provides "a detailed plan and timeline aiming for completion by October 2025." This meets the disclosure requirement when no policy exists under {{ESRS G1}} 10(d).

Figure 16 - Évaluation du niveau de conformité de la section G1-1 avec les exigences de l'ESRS.

Source : capture d'écran personnelle (juillet 2025), réalisée sur la plateforme NexusGPT.

Bien que l'agent auditeur permette de structurer une évaluation systématique de la conformité, cette démarche se heurte à une limite des LLM : les textes générés ne sont jamais strictement identiques d'une exécution à l'autre. Cela rend difficile une vérification totalement automatisée. Toutefois, l'outil peut aider les entreprises à identifier les manquements dans leurs réponses et à apporter plus de précision aux sections incomplètes.

Une autre réflexion importante a émergé concernant la cohérence d'ensemble du rapport, dans un contexte où chaque sous-section est générée indépendamment par un agent distinct. Plusieurs dispositifs ont été mis en place pour répondre à cet enjeu.

Tout d'abord, une fois les sections générées, la personne en charge de l'élaboration du rapport joue un rôle central dans la relecture, l'assemblage et la structuration des contenus. Elle veille à l'unité éditoriale, à la cohérence du style et à l'articulation fluide des différentes parties.

En parallèle, le client final est également impliqué : il peut relire les textes, proposer des ajustements ou compléter certains éléments afin de garantir que le rapport reflète fidèlement sa stratégie, ses priorités et ses données internes. Ces 2 niveaux d'intervention sont essentiels pour garantir la qualité du rapport, sa pertinence stratégique et son appropriation par l'entreprise.

Par ailleurs, le système repose sur une conception par sous-sections indépendantes, chaque GPT Task étant associée à une exigence précise des normes ESRS. Ce découpage fin permet de générer uniquement les parties pertinentes pour l'entreprise, en fonction de son cœur d'activité, de ses enjeux matériels et des résultats de son analyse de double matérialité. Ce fonctionnement offre un double avantage : éviter la production de contenu inutile et adapter finement le rapport à la réalité de chaque organisation. Il s'inscrit dans la logique même de la directive CSRD, qui impose une réponse ciblée, fondée sur la matérialité, et non une couverture exhaustive.

Enfin, la standardisation de la structure des prompts contribue également à l'homogénéité du rapport. Tous les agents respectent une architecture commune (ton, format, style narratif, consignes), ce qui garantit une cohérence rédactionnelle entre les sections, même lorsqu'elles sont produites de manière autonome.

7. Difficultés rencontrées lors du développement de l'agent

— Appropriation des concepts liés à la CSRD et aux ESRS

Avant de rejoindre le projet, je ne possédais aucune connaissance préalable de la directive CSRD. Il m'a donc fallu m'adapter rapidement et monter en compétences sur ce sujet. Cet apprentissage s'est avéré complexe, notamment en raison de la structure dense et peu intuitive des ESRS. Comme mentionné précédemment, l'architecture proposée par l'EFRAG rend la navigation entre les sections particulièrement laborieuse. Assimiler le vocabulaire technique ainsi que les nombreux acronymes a nécessité un temps d'adaptation conséquent.

— Montée en compétences sur la plateforme NexusGPT

Au moment du lancement du développement de l'outil, la plateforme NexusGPT était encore en cours de construction. Aucun guide n'était alors disponible pour accompagner la création d'agents, ce qui a engendré plusieurs difficultés et retardé l'avancement de la mission. L'apprentissage a donc reposé sur une autoformation exigeante, mobilisant un temps considérable. Ce n'est que vers la fin de ma période de stage que des supports pédagogiques ont été mis à disposition. En outre, la plateforme, encore instable, nous a confrontés à de nombreux bugs techniques que nous n'étions pas en mesure de résoudre de manière autonome, bloquant parfois l'évolution du projet. Ce manque d'autonomie nous rendait fortement dépendants des réunions hebdomadaires du jeudi avec le cofondateur, ce qui nous obligeait parfois à attendre plusieurs jours pour débloquer certaines situations.

— Construction et itération des prompts

La rédaction d'un prompt efficace s'est révélée particulièrement exigeante. Il m'a fallu trouver un équilibre entre la clarté des consignes et la précision des exigences posées par les ESRS.

Cette démarche de vulgarisation ne devait en aucun cas dénaturer le fond. Par ailleurs, bien que les agents produisent des réponses formellement bien structurées, le contenu manquait souvent de pertinence : de nombreuses phrases étaient inutiles ou hors sujet. Il a donc été nécessaire d'ajuster à plusieurs reprises les règles de comportement afin d'éviter la génération de contenu superflu ou de données inventées.

— **Collecte de données**

La phase de test de l'outil a nécessité une collecte de données qui s'est révélée problématique à plusieurs reprises. Si l'utilisation de ChatGPT pour générer des données fictives s'est avérée utile, ces dernières ne reflétaient pas la complexité d'une base de données réelle en entreprise. Les formes et les volumes de données varient fortement d'une organisation à l'autre, ce qui limite la portée des tests menés avec des jeux de données standardisés. Nous sommes toutefois parvenus à collecter certaines politiques internes issues d'Intys grâce à l'aide de consultants. Néanmoins, l'efficacité réelle de l'outil ne pourra être pleinement mesurée qu'à partir d'un cas client concret.

8. Contraintes techniques de l'outil et limites actuelles

8.1. Contraintes liées à l'IA et aux LLM (cadre théorique appliqué)

— **Les biais**

Comme évoqué dans les sections précédentes sur les limites de l'intelligence artificielle, des modèles LLM et des agents, la question des biais introduits lors de l'entraînement des modèles concerne également les agents que j'ai développés, puisqu'ils reposent eux-mêmes sur des modèles de langage pré-entraînés. Ces biais – qu'ils soient culturels, genrés ou contextuels — peuvent potentiellement influencer les contenus générés, bien que je n'aie pas pu les mesurer concrètement dans le cadre de ce projet.

— **Les hallucinations**

Les hallucinations, identifiées comme une des limites majeures des LLM et des agents IA, constituent également une contrainte pour les agents développés dans le cadre de ce travail. Comme observé lors de la phase de test, les agents peuvent parfois produire des phrases dénuées de sens ou contenant des informations incorrectes. Ce phénomène, déjà théorisé dans la partie sur les risques liés à l'IA générative, confirme la nécessité d'une validation humaine systématique.

— **La dépendance à de larges quantités de données**

Bien que je n'aie pas participé à la phase d'entraînement des modèles sous-jacents à NexusGPT, il n'en demeure pas moins que les agents que j'ai conçus reposent sur ces modèles, et héritent donc de leurs caractéristiques. Comme mentionné dans la section sur les limites de l'intelligence artificielle, ces modèles nécessitent d'importants volumes de données pour atteindre un bon niveau de performance, ce qui soulève des questions en termes d'énergie requise et d'impact environnemental.

— **Le prompt et la relecture humaine.**

Comme évoqué dans la section sur les considérations éthiques dans la conception de prompts, un prompt mal formulé – ambigu, incomplet ou biaisé – peut générer des réponses erronées ou incohérentes, accentuant ainsi le risque d'hallucination. Ce constat s'est vérifié dans la pratique, et souligne une limite importante : le niveau de confiance que l'on peut accorder aux réponses produites reste conditionné à une relecture attentive par un expert humain. Dans le cadre de la rédaction de rapports réglementaires, cette étape est indispensable.

— **La sécurité des données**

La question de la sécurité des données, soulevée à plusieurs reprises dans les sections consacrées aux risques liés aux LLM, aux prompts et aux agents, reste partiellement sans réponse dans le cadre de ce projet. S'agissant du partenariat entre Intys et NexusGPT, on m'a indiqué que les aspects liés à la sécurité étaient couverts contractuellement, mais je n'ai pas eu accès aux documents ni aux détails techniques correspondants. Par ailleurs, aucune information précise ne m'a été communiquée concernant les modalités de protection des données dans le cadre d'un futur contrat avec un client externe. Cette incertitude constitue une limite notable dans l'évaluation des risques du projet.

8.2. Contraintes techniques spécifiques au projet

— **Limitation du volume de texte traité**

L'une des principales contraintes techniques, qui a nécessité de nombreux ajustements dans la conception de l'agent, concerne la limite de caractères qu'une GPT Task peut traiter simultanément. La capacité maximale, d'environ 1 500 mots, a fortement complexifié le découpage et l'implémentation des standards ESRS au sein de l'outil.

— **Fiabilité de l'extraction d'information**

Une autre limite importante réside dans la fiabilité des informations extraites par l'agent à partir des documents sources. Une incertitude persiste quant à sa capacité à identifier l'ensemble des données exigées. Le contenu généré ne peut donc être considéré comme totalement fiable, rendant indispensable une relecture humaine. Par ailleurs, l'agent ne reproduit pas systématiquement une tâche de manière identique : 2 exécutions peuvent aboutir à des textes

sensiblement différents, ce qui complique l'évaluation de leur conformité de manière automatisée.

— **Instabilité de la plateforme**

Comme mentionné précédemment, le caractère encore en développement de la plateforme a engendré plusieurs difficultés techniques, notamment des bugs ou des interruptions liées à des opérations de maintenance. Ces instabilités ont freiné l'utilisation continue de l'outil et ralenti le rythme d'avancement du projet.

— **Reprise du projet en interne**

La brièveté de mon stage (60 jours) a soulevé des enjeux de transmission. Ayant été en grande partie seule à développer l'agent, il s'est avéré difficile pour une autre personne de s'approprier rapidement le fonctionnement de la plateforme et de l'outil conçu. L'entreprise devra ainsi former un nouveau collaborateur, à la fois à l'utilisation de NexusGPT et aux spécificités de l'agent développé. Cette reprise, décidée en interne, nécessitera une montée en compétences afin d'assurer la continuité du projet et de répondre aux besoins futurs des clients.

— **Impact du projet Omnibus et des outils déjà existants**

L'annonce du pack "Omnibus" par la Commission européenne – comme mentionné dans la section « Cadre initial de la CSRD » – a partiellement remis en question la pertinence initiale de l'outil. La réduction de près de 80 % du périmètre d'application limite considérablement le nombre d'entreprises concernées, fragilisant ainsi la portée potentielle de la solution développée.

— **Concurrence des outils existants sur le marché**

Au cours des entretiens menés dans le cadre de ce mémoire, 2 outils se sont particulièrement distingués en matière d'application de l'intelligence artificielle au service du reporting CSRD : *Karomia* et *Greenomy*.

Au moment du démarrage du projet, ces plateformes n'avaient pas encore été identifiées. Leur découverte tardive a mis en évidence l'existence d'une concurrence déjà bien établie, à laquelle l'outil développé devra nécessairement se confronter. Leur présence sur le marché confirme l'intérêt croissant pour l'usage de l'IA dans le processus de reporting. Déjà adoptées par de grandes entreprises, ces solutions bénéficient d'une légitimité renforcée par leur antériorité et leur facilité d'intégration, ce qui pourrait constituer un frein à l'adoption d'un outil interne, en particulier s'il n'est pas aussi mature ou reconnu.

Tombal (2025) revient sur ce choix d'outils :

« *Alors, dans un premier temps, on a assez vite réduit – on a exploré beaucoup d'outils – et on*

a assez vite réduit la liste à deux outils qui étaient Greenomy et Karomia. Pourquoi ces deux outils-là ? Parce que ce sont des outils qui permettent assez facilement de se mettre en ordre de marche, des outils qui sont assez faciles à prendre en main, où finalement on uploadé des documents et puis derrière, ça génère des réponses. »

Karomia

Karomia est une plateforme belge fondée en 2024 et basée à Gand. Spécialisée dans la création de rapports ESG, elle s'appuie sur l'intelligence artificielle pour générer des rapports conformes aux exigences de la CSRD ainsi qu'au standard volontaire VSME. Elle prend également en charge l'analyse de double matérialité, exigée par la directive (Karomia, n.d.-a).

Le processus repose sur une interface simple : l'utilisateur téléverse les documents existants de l'entreprise, quel qu'en soit le format. La plateforme procède ensuite à une analyse automatique visant à détecter les lacunes par rapport aux exigences des ESRS, que l'utilisateur est ensuite invité à compléter. Une fois cette étape réalisée, Karomia génère un rapport structuré conforme aux normes applicables (Karomia, n.d.-b).

L'outil a notamment été adopté par Lineas, qui l'utilise pour automatiser son analyse de double matérialité, une tâche longue et complexe. Vanoppen (2025) décrit ce processus :

« Nous avons utilisé l'intelligence artificielle pour réaliser l'analyse de double matérialité. C'est Karomia qui nous aide pour cela. Il s'agit de téléverser une série de documents, et l'IA, connaissant les articles des ESRS, recherche des indices et effectue une mise en correspondance. Elle retourne une longue liste d'IROs. Ces IROs sont ensuite soumis à des entretiens sur la plateforme, qui calcule automatiquement la matrice DMA. C'est donc entièrement automatisé. Et elle génère un rapport DMA. J'ai un rapport de 70 pages qui a été entièrement rédigé par l'IA. »

Greenomy

Greenomy est une plateforme bruxelloise fondée en 2021, spécialisée dans le reporting en matière de durabilité. Elle utilise l'intelligence artificielle pour générer des rapports conformes aux principales réglementations européennes, telles que la CSRD, le standard volontaire VSME et la taxonomie (Greenomy, n.d.-b). Selon les documents de l'entreprise, la solution permettrait d'accélérer le reporting, le rendant jusqu'à 7 fois plus rapide qu'un traitement manuel (Greenomy, n.d.-a).

Les données sont centralisées dans une bibliothèque ESG et peuvent être enrichies, structurées et exploitées via une interface collaborative. Greenomy propose également des services de conseil complémentaires : analyse de double matérialité, évaluation de conformité (gap assessment) et relecture avant audit externe (Greenomy, n.d.-a).

Tombal (2025) partage son point de vue sur le choix des outils retenus par son organisation :

« Greenomy a une manière de fonctionner très décentralisée, où il faut vraiment aller data point par data point, solliciter les contributeurs qui vont aller mettre l'information, uploader le document pour ce data point spécifique, générer une réponse, la corriger directement. Donc c'est un travail qui est très procédurier, et qui est très bien dans une organisation où les gens peuvent avoir énormément de temps à y consacrer, et sont réactifs par rapport à ce genre de demande. Là où Karomia a plus une vue, entre guillemets, centralisée, où on peut avoir une personne qui centralise l'ensemble des documents, upload le tout en batch, et puis après ça se gère et se dispatche, mais à partir de quelque chose qui est déjà construit. Et ça correspondait plus à notre manière de travailler ici chez Sibelga, parce qu'on savait que de toute façon, pour que le projet CSRD avance, il faudrait qu'il y ait un leadership fort, avec quelqu'un qui centralise et qui pousse un petit peu. Et ça nous semblait plus facile de faire ça avec Karomia qu'avec Greenomy. »

9. Élargissement du champ d'application : vers le reporting ESG

La présence de solutions concurrentes déjà bien établies, conjuguée à la réduction significative du champ d'application de la CSRD introduite par le paquet Omnibus, a conduit à une réorientation stratégique du projet. Plutôt que de poursuivre uniquement avec les agents initialement conçus pour la seule conformité à la CSRD, une nouvelle génération d'agents a été développée dans une optique plus large.

Ces nouveaux agents sont conçus pour générer des rapports ESG plus généralistes, mieux adaptés à une diversité de cas d'usage et de profils d'entreprises – y compris celles non soumises à une obligation réglementaire. Cette évolution reflète à la fois la maturité croissante du marché et l'émergence d'une demande structurée pour des outils automatisés de reporting en matière de durabilité.

10. Retour d'expérience

D'un point de vue personnel et intellectuel, ce projet m'a permis de découvrir concrètement les défis liés à la conception d'un projet à fort enjeu, en autonomie. N'ayant jamais dirigé un projet seule auparavant, ce fut un véritable défi pour moi. L'envie de bien faire et de réussir a parfois généré une pression importante, que j'ai dû apprendre à gérer. Ce travail m'a ainsi aidée à prendre du recul, à mieux canaliser le stress, et à développer une forme de résilience, notamment face aux détours, aux retours en arrière ou aux doutes qui ont marqué le parcours. J'ai appris à ne pas me laisser emporter par l'enthousiasme des débuts, mais à adopter une posture plus méthodique et stable sur la durée.

Ce projet m'a également permis de développer une grande autonomie. En l'absence de mes responsables à certains moments, j'ai dû prendre des décisions seule, résoudre des problèmes

techniques, ou avancer malgré les incertitudes. Cela m'a poussée à prendre des initiatives, à aller chercher des solutions par moi-même, et à structurer ma démarche avec rigueur. J'ai aussi beaucoup progressé dans ma capacité à communiquer clairement mes idées, notamment en expliquant le fonctionnement d'un agent IA génératrice de rapports CSRD à des interlocuteurs non techniques. Ce travail m'a sensibilisée à un point essentiel : la clarté pédagogique. Ce n'est pas parce qu'on maîtrise un sujet que la personne en face le comprend, et il est donc crucial de savoir adapter son discours.

Sur le plan technique, j'ai énormément appris. Je ne connaissais pas la directive CSRD au départ, et j'ai fini par m'y plonger avec un véritable intérêt. L'analyse du texte, combinée aux retours très concrets obtenus grâce aux entretiens, m'a permis de comprendre les enjeux de mise en œuvre pour les entreprises, la diversité d'interprétation, et les implications concrètes. J'ai aussi significativement renforcé mes compétences dans l'usage de ChatGPT, que j'ai utilisé pour concevoir, tester et ajuster les prompts. Cette expérience m'a permis de mieux appréhender les limites et les leviers d'un agent génératif dans un cadre normatif exigeant. Ce qui m'a particulièrement surprise, c'est le niveau de finesse qu'un simple prompt peut exiger pour produire un résultat pertinent : un mot oublié, une tournure ambiguë, et le résultat obtenu change complètement.

Avec le recul, l'un des éléments que je ferais différemment serait de ne pas hésiter à m'exprimer davantage et à poser des questions dès le départ. Lorsqu'on aborde un domaine nouveau et que l'on interagit avec des interlocuteurs perçus comme experts, il est souvent instinctif de rester en retrait, d'écouter et de ne pas oser intervenir, par crainte de poser une « mauvaise » question ou de paraître peu légitime. Ce fut mon cas, notamment lors des premières réunions avec Assem, au début du projet. Cette retenue a sans doute ralenti ma courbe d'apprentissage et généré une certaine frustration inutile. Aujourd'hui, avec l'expérience acquise, je sais que poser des questions – même simples – est souvent la meilleure manière de progresser. Ce projet m'a permis de gagner en assurance, et je me sens désormais plus légitime et plus confiante à l'idée d'initier ou de piloter de futurs projets dans ma carrière professionnelle.

11. Analyse critique du projet et de ses choix méthodologiques

L'initiative de ce projet s'est révélée ambitieuse, mais globalement réussie sur le plan conceptuel. À l'issue de mon stage, j'ai pu concevoir 4 agents dédiés à la rédaction de rapports CSRD, 3 agents auditeurs ainsi que 3 autres orientés vers le reporting ESG au sens plus large. L'idée de recourir à des agents d'intelligence artificielle pour assister les entreprises dans la rédaction de leurs rapports de durabilité m'apparaît tout à fait pertinente. Cette tâche, souvent longue et rigoureuse, se prête bien à l'automatisation partielle par des outils génératifs. En ce sens, le projet avait une véritable légitimité.

Le choix de la plateforme NexusGPT s'inscrivait dans cette logique. Elle permet en effet de concevoir des agents sans nécessiter de compétences en programmation, tout en offrant une

grande marge de personnalisation. Pour une entreprise comme Intys, qui cherche à enrichir son offre de services, cette démarche constitue un positionnement stratégique intéressant.

Toutefois, plusieurs limites sont à souligner. D'un point de vue organisationnel, confier à une stagiaire la responsabilité d'un projet à fort enjeu – tant sur le plan technique que commercial – représente un pari risqué. Même si cette expérience a été extrêmement formatrice pour moi, elle soulève des questions de continuité. À la fin d'un stage, la personne ayant acquis l'expertise nécessaire quitte l'entreprise. Sans un transfert de compétences structuré ou une reprise rapide du projet par des collaborateurs internes, le risque est de devoir recommencer la formation à zéro, ce qui mobilise à nouveau du temps et des ressources.

Par ailleurs, la création d'agents efficaces sur NexusGPT nécessite des compétences techniques spécifiques, notamment en conception de prompts. Comme exposé dans les sections précédentes, la qualité des résultats générés dépend fortement de la formulation. Concevoir un prompt optimal exige rigueur, itérations, et compréhension fine du fonctionnement des modèles de langage.

Enfin, le choix de NexusGPT, bien que pertinent, s'est heurté à une contrainte de calendrier. La plateforme était encore en cours de développement durant la période de mon stage, ce qui a occasionné plusieurs interruptions, bugs ou changements liés aux mises à jour. Travailler avec un outil en phase de stabilisation technique a indéniablement ralenti le bon déroulement du projet.

Malgré la satisfaction d'avoir pu mener à bien ce projet dans les délais impartis, je ressors également avec un léger sentiment d'inachevé. N'ayant pas eu l'opportunité de prolonger moi-même l'expérimentation ni de participer activement à son déploiement, je reste dans l'incertitude quant à la suite qui lui sera donnée. Cette limite renforce ma prise de conscience de l'importance du passage de la phase théorique à la phase d'implémentation, souvent déterminante dans la réussite concrète d'un projet.

Conclusion

Dans un contexte où l'intelligence artificielle occupe une place croissante dans les pratiques professionnelles, ce mémoire a exploré son potentiel afin de répondre à un besoin concret : accompagner les entreprises dans la rédaction de leurs rapports de durabilité requis par la directive CSRD. L'objectif principal était de concevoir, à l'aide de la plateforme NexusGPT, un agent d'intelligence artificielle capable de générer ces rapports de manière structurée et conforme.

J'ai structuré mon travail autour de 4 volets complémentaires. Les 3 premiers, théoriques, m'ont permis de comprendre les fondements de la CSRD ainsi que les concepts clés liés à l'intelligence artificielle et aux agents génératifs. Le quatrième, pratique, a consisté à concevoir puis tester les outils reposant sur ces principes, en développant 4 agents spécialisés selon les grandes sections des standards ESRS.

Ce travail a révélé toute la complexité de la conception d'agents génératifs. Si la plateforme NexusGPT ne nécessite pas de compétences en programmation, elle impose en revanche une grande rigueur dans la structuration des prompts. J'ai découvert à quel point un mot mal formulé pouvait altérer la qualité des réponses générées. Aujourd'hui, les agents sont opérationnels, mais leur efficacité ne pourra être pleinement évaluée qu'à travers leur utilisation dans un contexte réel, avec des données d'entreprise authentiques.

Ce projet m'a également permis de prendre conscience des opportunités offertes par l'IA générative dans le domaine du reporting. Comme l'illustrent les concurrents identifiés (Karomia, Greenomy), le marché est en plein essor. Je suis convaincue qu'Intys a ici une carte à jouer pour se démarquer grâce à une offre innovante, personnalisable et accessible.

Enfin, ce projet m'a beaucoup apporté sur le plan personnel. J'ai développé mon autonomie, ma capacité d'analyse, ma rigueur méthodologique. J'ai appris à gérer un projet de bout en bout, à faire face aux imprévus, à assumer des responsabilités et à ajuster mes méthodes de travail. Autant d'expériences qui me serviront dans la suite de mon parcours professionnel, quel que soit le domaine que je choisirai.

Bibliographie

- Abro, A. A., Talpur, M. S. H., & Jumani, A. K. (2022). Natural Language Processing Challenges and Issues: A Literature Review. *Gazi University Journal of Science*. <https://doi.org/10.35378/gujs.1032517>
- Acquier, A., & Gond, J.-P. (2007). Aux sources de la responsabilité sociale de l'entreprise : à la (re)découverte d'un ouvrage fondateur, Social Responsibilities of the Businessman d'Howard Bowen. *ResearchGate*. https://www.researchgate.net/publication/4797228_Aux_sources_de_la_responsabilite_sociale_de_l%27entreprise_a_la_redecouverte_d%27un_ouvrage_fondateur_Social_Responsibilities_of_the_Businessman_d%27Howard_Bowen
- Adams, C., Alhamood, A., Wang, L., & Wang, Y. (2021). *The double-materiality concept Application and issues Invited contribution*. <https://www.globalreporting.org/media/jrbntbyv/griwhitepaper-publications.pdf>
- Afnor Groupe. (n.d.). *L'ISO 26000 en 10 questions*. <https://www.afnor.org/wp-content/uploads/2016/08/ISO26000-en-10-questions.pdf>
- Alabi, A., & Gbemigb, O. (2025, 9 mai). *The Evolution of NLP: A Journey Through Time World Wide And It Important*. Research Gate. https://www.researchgate.net/publication/391594579_The_Evolution_of_NLP_A_Journey_Through_Time_World_Wide_And_It_Important
- Attarça, M., & Jacquot, T. (2005). *La représentation de la Responsabilité Sociale des Entreprises : une confrontation entre les approches théoriques et les visions managériales*. AIMS. <https://www.strategie-aims.com/conferences/9-xiveme-conference-de-l-aims/communications/603-la-representation-de-la-responsabilite-sociale-des-entreprises-une-confrontation-entre-les-approches-theoriques-et-les-visions-manageriales/download>
- Aureli, S., Salvatori, F., & Magnaghi, E. (2020). A Country-Comparative Analysis of the Transposition of the EU Non-Financial Directive: An Institutional

- Approach. *Accounting, Economics, and Law: A Convivium*, 10(2).
- <https://doi.org/10.1515/ael-2018-0047>
- Ayming. (2024, March 13). *RSE et ESG : quelles différences ?* Ayming France.
- <https://www.ayming.fr/insights/actualites/rse-et-esg-quelles-differences/>
- Baumüller, J., & Grbenic, S. O. (2021). Moving from non-financial to sustainability reporting: Analyzing the EU Commission's proposal for a Corporate Sustainability Reporting Directive (CSRD). *Facta Universitatis, Series: Economics and Organization*, 18, 369–381.
- <http://casopisi.junis.ni.ac.rs/index.php/FUEconOrg/article/view/8149/4374>
- Belidan, M. A., & Baghad, H. (2024). Navigating materiality in sustainability reporting: A scoping review of single and double materiality approaches. *Zeszyty Teoretyczne Rachunkowości*, 48(4), 11–32. <https://doi.org/10.5604/01.3001.0054.8687>
- Benani, A. (2024). Histoire et perspectives de l'intelligence artificielle. *Médecine/Sciences*, 40(3), 283–286.
- <https://doi.org/10.1051/medsci/2024022>
- Billio, M., Costola, M., Hristova, I., Latino, C., & Pelizzon, L. (2021). Inside the ESG ratings: (Dis)agreement and Performance. *Corporate Social Responsibility and Environmental Management*, 28(5), 1426–1445. <https://doi.org/10.1002/csr.2177>
- Boyer-Allirol, B. (2023, mai). Faut-il mieux réglementer le reporting extra-financier pour améliorer sa prise en compte par les investisseurs ? Comptabilité sans Frontières – The French Connection (Canada). <https://hal.science/hal-00992961v1/document>
- Briouya, A., Briouya, H., & Choukri, A. (2024). Overview of the progression of state-of-the-art language models. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 22(4), 897–909.
- <https://telkomnika.uad.ac.id/index.php/TELKOMNIKA/article/view/25936/11925>
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., & Hesse, C.

- (2020). *Language Models are Few-Shot Learners*. arXiv.
<https://arxiv.org/pdf/2005.14165>
- Burke, S. A., & Akhtar , A. (2023). The shortcomings of artificial intelligence: A comprehensive study. *International Journal of Library and Information Science*, 15(2), 8–13.
https://www.researchgate.net/publication/373536152_The_shortcomings_of_artificial_intelligence_A_comprehensive_study
- Capron, M., & Quairel, F. (2003). REPORTING SOCIETAL : LIMITES ET ENJEUX DE LA PROPOSITION DE NORMALISATION INTERNATIONALE " GLOBAL REPORTING INITIATIVE ". <https://shs.hal.science/halshs-00582742>
- Carroll, A. B. (1979). A Three-Dimensional Conceptual Model of Corporate Performance. *Academy of Management Review*, 4(4), 497–505.
https://www.researchgate.net/publication/230745468_A_Three-Dimensional_Conceptual_Model_of_Social_Performance
- Chai, Y., Jin, L., Feng, S., & Xin, Z. (2024). Evolution and advancements in deep learning models for Natural Language Processing. *Applied and Computational Engineering*, 77(1), 144–149. <https://doi.org/10.54254/2755-2721/77/20240674>
- Commission européenne. (2001). *Livre vert : Promouvoir un cadre européen pour la responsabilité sociale des entreprises: Vols. COM(2001) 366 final.*
[https://www.europarl.europa.eu/meetdocs/committees/deve/20020122/com\(2001\)366_fr.pdf](https://www.europarl.europa.eu/meetdocs/committees/deve/20020122/com(2001)366_fr.pdf)
- Commission européenne. (2011). Responsabilité sociale des entreprises : une nouvelle stratégie de l'Union européenne pour 2011-2014. <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0681:FIN:fr:PDF>
- Commission européenne. (2021, avril 21). *Proposition de directive du Parlement européen et du Conseil modifiant la directive 2013/34/UE, la directive 2004/109/CE, la directive 2006/43/CE et le règlement (UE) n° 537/2014, en ce qui concerne la publication d'informations en matière de durabilité par les entreprises*. <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:52021PC0189>

Commission européenne. (2025, février 26). *Questions\réponses sur le train de mesures “omnibus”*. https://ec.europa.eu/commission/presscorner/detail/fr/qanda_25_615

Connaissance des Énergies. (2021, août 25). *Marée noire : quelles ont été les plus importantes au cours des dernières décennies ?* Connaissance Des Énergies. <https://www.connaissancedesenergies.org/questions-et-reponses-energies/transport-maritime-quelles-sont-les-plus-grandes-marees-noires-des-dernieres-decennies>

Daga, D.-S. (2025, 10 avril). *Sustainability & CSR Reporting Manager chez Alstom* [Entretien]. Bruxelles.

Decaux, L. (2025, 4 août). *Head of Managerial Research, Development & Innovation chez Intys* [Entretien]. Bruxelles.

Deshpande, S.-A. (2025, 2 avril). *ESG Expert chez Elia* [Entretien]. Bruxelles.

EFRAG. (n.d.). *ESRS implementation guidance documents*. Efrag.org. <https://www.efrag.org/en/projects/esrs-implementation-guidance-documents>

EFRAG. (2023a). *ESRS Set 1*. Efrag.org. <https://xbrl.efrag.org/e-esrs/esrs-set1-2023.html>

EFRAG. (2023b, juillet 31). *EFRAG welcomes the adoption of the Delegated Act on the first set of European Sustainability Reporting Standards (ESRS) by the European Commission*. EFRAG. https://www.efrag.org/en/news-and-calendar/news/efrag-welcomes-the-adoption-of-the-delegated-act-on-the-first-set-of-european-sustainability?utm_source=chatgpt.com

EFRAG. (2023c, novembre 30). *EFRAG and GRI enhance collaboration with deeper ties*. EFRAG. <https://www.efrag.org/en/news-and-calendar/news/efrag-and-gri-enhance-collaboration-with-deeper-ties>

EFRAG. (2024). *EFRAG IG 2 :Value Chain Implementation Guidance*. https://www.efrag.org/sites/default/files/sites/webpublishing/SiteAssets/EFRAG%20IG%202%20Value%20Chain_final.pdf

Elia Group. (2024). Securing Clean Competitiveness - Rapport Annuel Intégré 2024 (p. 157). Elia Group. <https://investor.eliagroup.eu/en/reports-and-results/reports-for-elia-group>

Encyclopédie d'histoire numérique de l'Europe. (2025). *10 juillet 1976 : la catastrophe de Seveso*. Encyclopédie d'Histoire Numérique de L'Europe.

- <https://ehne.fr/fr/encyclopedie/th%C3%A9matiques/écologies-et-environnements/les-risques-environnementaux/10-juillet-1976-la-catastrophe-de-seveso>
- EUR-Lex. (2022, janvier 7). *Règlement général sur la protection des données (RGPD)*.
<https://eur-lex.europa.eu/FR/legal-content/summary/general-data-protection-regulation-gdpr.html>
- European Commission. (n.d.). *ESG Rating Activities*. Finance.
https://finance.ec.europa.eu/sustainable-finance/tools-and-standards/esg-rating-activities_en
- European Commission. (2024). *Corporate Sustainability Reporting*. European Commission.
https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en
- European Parliament. (2023, August 6). *EU AI Act: first regulation on artificial intelligence | Topics | European Parliament*. Topics | European Parliament.
<https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence#ai-regulation-in-europe-the-first-comprehensive-framework-4>
- EY. (2024). *Issue 4 / EU Sustainability Developments*. <https://www.ey.com/content/dam/ey-unified-site/ey-com/en-gl/technical/ifrs-technical-resources/documents/ey-eu-sustainability-dev-4-february-2024.pdf>
- Fairfax, L. (2023). ESG hypocrisy and voluntary disclosure. *NYU Journal of Legislation and Public Policy*, 26(127).
https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=1441&context=faculty_articles
- Fornasari, T., & Traversi, M. (2024, juillet 11). *The Impact of the CSRD and the ESRS on Non-Financial Disclosure*. ResearchGate; Niccolo Cusano University-Rome Symphonya Emerging Issues in Management.
https://www.researchgate.net/publication/382206526_The_Impact_of_the_CSRD_and_the_ESRS_on_Non-Financial_Disclosure

- Friedman, M. (1970, septembre 13). The Social Responsibility of Business Is to Increase Its Profits. *The New York Times*. <https://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html>
- Gorlin, C. (2023, janvier 19). *The European Sustainability Reporting Standards : a Phase-in Journey to Sustainability Alignment*. *Greenconomy.io*.
<https://www.greenconomy.io/blog/esrs-phase-in-journey>
- Greenconomy. (n.d.-a). *Book a demo*. Greenconomy. https://www.greenconomy.io/book-a-demo?utm_source=brochure+&utm_medium=website&utm_campaign=Website+Brochure+Downloads&utm_id=brochureen
- Greenconomy. (n.d.-b). *À propos*. LinkedIn.
<https://www.linkedin.com/company/greenomysaas/about/>
- GRI. (n.d.). *Mission & history*. <https://www.globalreporting.org/about-gri/mission-history/>
- Gupta, P., Ding, B., Guan, C., & Ding, D. (2024). Generative AI: A systematic review using topic modelling techniques. *Data and Information Management*, 8(2), 100066.
<https://doi.org/10.1016/j.dim.2024.100066>
- Gupta, R. K. (2025, janvier 18). *Limitations and Future Role of AI Systems*. Research Gate.
<https://doi.org/10.13140/RG.2.2.14270.60484>
- Hales, J. (2021). Sustainability Accounting Standards Board (SASB). *World Scientific Encyclopedia of Climate Change*, 3, 37–41.
https://www.worldscientific.com/doi/epdf/10.1142/9789811213960_0007
- Hettiarachchi, I. (2025). Exploring Generative AI Agents: Architecture, Applications, and Challenges. *Journal of Artificial Intelligence General Science (JAIGS) ISSN:3006-4023*, 8(1), 105–127. <https://doi.org/10.60087/jaigs.v8i1.350>
- Hirani, R., Noruzi, K., Khuram, H., Hussaini, A. S., Aifuwa, E. I., Ely, K. E., Lewis, J. M., Gabr, A. E., Smiley, A., Tiwari, R. K., & Etienne, M. (2024). Artificial Intelligence and Healthcare: A Journey through History, Present Innovations, and Future Possibilities. *Life*, 14(5), 557. <https://doi.org/10.3390/life14050557>
- Horne, D. (2025). The Agentic AI Mindset -A Practitioner's Guide to Architectures, Patterns, and Future Directions for Autonomy and Automation. ResearchGate.

- https://www.researchgate.net/publication/390958865_The_Agentic_AI_Mindset_-A_Practitioner's_Guide_to_Architectures_Patterns_and_Future_Directions_for_Autonomy_and_Automation.
- IBM. (2021a, March 2). *Format JSON (JavaScript Object Notation)*. IBM.
<https://www.ibm.com/docs/fr/bpm/8.5.6?topic=formats-javascript-object-notation-json-format>
- Idowu, M., Cena, J., & Godwin Olaoye. (2024, novembre 13). *Exploring the limitations and challenges of GPT models*. ResearchGate.
https://www.researchgate.net/publication/385771524_EXPLORING_THE_LIMITATIONS_AND_CHALLENGES_OF_GPT_MODELS
- IFRS. (2024, mai 2). *IFRS Foundation and EFRAG publish interoperability guidance*. IFRS.
<https://www.ifrs.org/news-and-events/news/2024/05/ifrs-foundation-and-efrag-publish-interoperability-guidance/>
- Inês Hipólito. (2024, août 7). *The Context Windows Fallacy in Large Language Models*. Research Gate.
https://www.researchgate.net/publication/382941928_The_Context_Windows_Fallacy_in_Large_Language_Models
- Intys. (2023). *Intys 2023 CSR commitments & approach*. [PDF]. Intys.
<https://www.intys.eu/engage-for-our-consultants>
- intys. (n.d.). *Intys*. <https://www.intys.eu>
- Intysify| Operational Sustainability Consulting. (2023). *Intysify| Operational Sustainability Consulting*. Intysify| Operational Sustainability Consulting. <https://www.intysify.eu>
- ISO. (n.d.). *ISO - ISO 26000 – Responsabilité sociétale*. ISO. <https://www.iso.org/fr/iso-26000-social-responsibility.html>
- ISO. (2022). *Norme ISO/IEC 27001 – Systèmes de management de la sécurité de l'information*. ISO. <https://www.iso.org/fr/standard/27001>
- ISO. (2023). *ISO/IEC 42001:2023*. ISO. <https://www.iso.org/fr/standard/42001>
- Jamin, O. (2025, 15 avril). *Sustainability Manager chez Veolia* [Entretien]. Bruxelles.

- Jbara, N. (2017). Perspective historique de la responsabilité sociale des entreprises (RSE). *Revue Multidisciplinaire Sur l'Emploi, Le Syndicalisme et Le Travail*, 11(1), 86–102. <https://doi.org/10.7202/1043839ar>
- Johnson, S., & Hyland-Wood, D. (2025). A Primer on Large Language Models and their Limitations. Qeios. <https://doi.org/10.32388/nhjyvs>
- Kain, T. (2025, mars 18). *Responsabilité sociale des entreprises et ESG : comment elles fonctionnent ensemble*. Sigma Earth.
<https://sigmaearth.com/fr/responsabilite%C3%A9-de-d%C3%A9veloppement-durable-des-entreprises-et-ESG-%3A-comment-ils-fonctionnent-ensemble/>
- Kalyan, K. S. (2023). A Survey of GPT-3 Family Large Language Models Including ChatGPT and GPT-4. *ResearchGate*. <https://doi.org/10.13140/RG.2.2.18555.49441>
- Kapoor, S., Stroebel, B., Siegel, Z., Nadgir, N., & Narayanan, A. (2024). *AI Agents That Matter*. <https://arxiv.org/pdf/2407.01502>
- Karomia. (n.d.-a). *À propos*. LinkedIn. <https://www.linkedin.com/company/karomia/about/>
- Karomia. (n.d.-b). *AI CSRD Reporting Tool for EU Compliance*. Karomia.
<https://www.karomia.eu/solutions/csr>
- Kaźmierczak, M. (2022). A literature review on the difference between CSR and ESG. *Scientific Papers of Silesian University of Technology. Organization and Management Series*, 2022(162), 275–289. <https://doi.org/10.29119/1641-3466.2022.162.16>
- Keaveny, E. (2024). *CSRD and double materiality: where to begin*. EY.
https://www.ey.com/en_ie/insights/sustainability/csr-how-to-manage-your-double-materiality-assessment
- Koutaya, A., & Koutaya, Y. (2020, May 11). *Essai de synthèse des principales étapes d'évolution de la responsabilité sociétale des entreprises*. *ResearchGate*.
<https://www.researchgate.net/publication/341293777>
- KPMG. (n.d.). *Ready for the change? An Analysis of Irish Companies' ESG Reporting Readiness*. <https://assets.kpmg.com/content/dam/kpmg/ie/pdf/2021/12/ie-csr-readiness-study.pdf>

- KPMG. (2022a). Big shifts, small steps Survey of Sustainability Reporting 2022.
<https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2022/10/ssr-small-steps-big-shifts.pdf>
- KPMG. (2022b). The Corporate Sustainability Reporting Directive (CSRD).
<https://assets.kpmg.com/content/dam/kpmgsites/ch/pdf/the-corporate-sustainability-reporting-directive.pdf>
- KPMG. (2024a). *Get ready for the Corporate Sustainability Reporting Directive Understanding the CSRD.*
<https://assets.kpmg.com/content/dam/kpmg/nl/pdf/2024/services/faq-csrd-2024.pdf>
- KPMG. (2024b). The move to mandatory reporting Survey of Sustainability Reporting 2024.
<https://assets.kpmg.com/content/dam/kpmgsites/xx/pdf/2024/11/the-move-to-mandatory-reporting-web-copy.pdf.coredownload.inline.pdf>
- KPMG Ireland. (2019). *The numbers that are changing the world Revealing the growing appetite for responsible investing.*
<https://assets.kpmg.com/content/dam/kpmg/ie/pdf/2019/10/ie-numbers-that-are-changing-the-world.pdf>
- Krishnan, N. (2025). AI Agents: Evolution, Architecture, and Real-World Applications. Arxiv. <https://arxiv.org/pdf/2503.12687>
- Levy, M., Jacoby, A., & Goldberg, Y. (2024). *Same Task, More Tokens: the Impact of Input Length on the Reasoning Performance of Large Language Models.* ArXiv.
<https://arxiv.org/pdf/2402.14848>
- Liu, N. F., Lin, K., Hewitt, J., Paranjape, A., Bevilacqua, M., Petroni, F., & Liang, P. (2024). Lost in the Middle: How Language Models Use Long Contexts. *Transactions of the Association for Computational Linguistics, 12*, 157–173.
https://doi.org/10.1162/tacl_a_00638
- Mahesh, B. (2018). Machine learning algorithms - a review. *International Journal of Science and Research (IJSR) ResearchGate Impact Factor, 9*(1).
<https://doi.org/10.21275/ART20203995>
- Mangold, R. (2025, 23 avril). *Responsable RSE chez T&S* [Entretien]. Bruxelles.

- Mezzanotte, F. E. (2023). Corporate sustainability reporting: double materiality, impacts, and legal risk. *Journal of Corporate Law Studies*, 23(2), 1–31.
<https://doi.org/10.1080/14735970.2024.2319058>
- Moradi, M., Yan, K., Colwell, D., Samwald, M., & Asgari, R. (2025). A Critical Review of Methods and Challenges in Large Language Models. *Computers, Materials and Continua*, 82(2), 1681–1698.
<https://www.sciencedirect.com/org/science/article/pii/S1546221825000992>
- Mukile, P. (2025, 17 mars). *Head of Audit, Risk & Quality chez Sibelga* [Entretien]. Bruxelles.
- Mulkens, S. (2024, décembre 6). *The move to mandatory reporting*. KPMG.
<https://kpmg.com/be/en/home/insights/2024/12/sus-the-move-to-mandatory-reporting.html>
- Mullenbach-Servayre, A. (2007). L'apport de la théorie des parties prenantes à la modélisation de la responsabilité sociétale des entreprises. *La Revue Des Sciences de Gestion*, 223(1), 109. <https://doi.org/10.3917/rsg.223.0109>
- NexusGPT. (2024). *NexusGPT*. <https://gpt.nexus>
- NexusGPT. (2025a). *Adding Knowledge to Your Agent*. <https://learn.gpt.nexus/key-features/adding-knowledge-to-your-agent>
- NexusGPT. (2025b). *Adding Skills to Your Agent*. <https://learn.gpt.nexus/key-features/adding-skills-to-your-agent>
- NexusGPT. (2025c). *Creating Your Agent*. <https://learn.gpt.nexus/key-features/creating-your-agent>
- NexusGPT. (2025d). *Foundational Knowledge*. <https://learn.gpt.nexus/key-concepts>
- NexusGPT. (2025e). *Updating Your Agent's Prompt*. <https://learn.gpt.nexus/key-features/updating-your-agents-prompt>
- Nielsen, C. (2023). ESG Reporting and Metrics: From Double Materiality to Key Performance Indicators. *Sustainability*, 15(24), 16844.
<https://doi.org/10.3390/su152416844>

- Novethic. (2025, janvier 26). *Catastrophe de Bhopal : causes et conséquences* - Novethic. Novethic. <https://www.novethic.fr/lexique/detail/bhopal.html>
- OCDE. (2023). *Principes directeurs de l'OCDE à l'intention des entreprises multinationales sur la conduite responsable des entreprises*. OCDE. <https://doi.org/10.1787/0e8d35b5-fr>
- OpenAI. (n.d.-a). Models- *OpenAI API*. OpenAI. <https://platform.openai.com/docs/models>
- OpenAI. (n.d.-b). *What is ChatGPT? | OpenAI Help Center*. OpenAI. <https://help.openai.com/en/articles/6783457-what-is-chatgpt>
- OpenAI. (2025, février 27). *Introducing GPT-4.5*. OpenAI. <https://openai.com/index/introducing-gpt-4-5/>
- Organisation internationale du Travail. (n.d.). *Déclaration de principes tripartite concernant les entreprises multinationales et la politique sociale (Déclaration sur les EMN)*. Organisation internationale du Travail. <https://www.ilo.org/fr/departement-des-entreprises/domaines-de-travail/declaration-de-principes-tripartite-concernant-les-entreprises>
- Pacte Mondial Réseau France. (n.d.). *Pacte mondial des Nations Unies & pacte mondial ONU*. Pacte Mondial Réseau France. <https://pactemondial.org/decouvrir/pacte-mondial-des-nations-unies/>
- Pantazi, T. (2024). The Introduction of Mandatory Corporate Sustainability Reporting in the EU and the Question of Enforcement. *European Business Organization Law Review*, 25. <https://doi.org/10.1007/s40804-024-00320-x>
- Passas, I. (2024). The Evolution of ESG: From CSR to ESG 2.0. *Encyclopedia*, 4(4), 1711–1720. <https://doi.org/10.3390/encyclopedia4040112>
- PwC. (n.d.-a). *CSRD - Double Materiality Assessment*. PwC. <https://www.pwc.nl/en/topics/sustainability/esg/corporate-sustainability-reporting-directive/csrn-double-materiality-assessment.html>

- PwC. (n.d.-b). *The Corporate Sustainability Reporting Directive (CSRD)*. PwC. <https://www.pwc.com/lv/en/about/services/sustainability-services/the-corporate-sustainability-reporting-directive-CSRD.html>
- PwC. (2024, janvier 31). *Understanding the CSRD Double Materiality Assessment Process*. PwC. <https://www.pwc.com/mt/en/publications/sustainability/understanding-the-csrd-double-materiality-assessment-process.html>
- Raghavan, K. (2022). ESG Reporting Impact on Accounting, Finance. *The Journal of Global Awareness*, 3(1), 1–16. <https://doi.org/10.24073/jga/3/01/09>
- Ramos, M. C., Collison, C. J., & White, A. D. (2024). A Review of Large Language Models and Autonomous Agents in Chemistry. *Chemical Science*, 16(6). <https://pubs.rsc.org/en/content/articlelanding/2025/sc/d4sc03921a>
- République française. (2024, octobre 15). *Double matérialité : l'analyse des impacts, risques et opportunités*. Portail RSE. <https://portail-rse.beta.gouv.fr/csrn,double-materialite-comprendre-et-appliquer-cette-analyse-essentielle-pour-le-reporting-de-durabilite/>
- République française. (2025, février 27). *Seuils CSRD et projet de directive Omnibus*. Portail RSE. <https://portail-rse.beta.gouv.fr/csrn/seuils-csrn-omnibus-criteres-d-application/>
- Sauoussany, A., & Kidaye, N. (2023). Les Déterminants du Reporting RSE : Etude de l'Impact des Mécanismes de Gouvernance sur la Diffusion Sociétale. *Research Gate*. https://www.researchgate.net/publication/371695494_Les_Determinants_du_Reportin g_RSE_Etude_de_l'Impact_des_Mecanismes_de_Gouvernance_sur_la_Diffusion_Soc ietale
- SASB. (2023, juin 20). *About us*. SASB. <https://sasb.ifrs.org/about/>
- Scarcello, F. (2019). Artificial Intelligence. *Elsevier EBooks*, 287–293. <https://www.sciencedirect.com/topics/neuroscience/natural-language-processing>
- SecureFrame. (2025). *SOC 2 Overview | Secureframe*. Secureframe. <https://secureframe.com/hub/soc-2/overview#beginner>
- Shannon, M. (2024, février 28). *Limited vs reasonable assurance over ESG*. KPMG. <https://kpmg.com/xx/en/our-insights/esg/limited-vs-reasonable-assurance-over-esg.html>

Sharifani, K., & Amini, M. (2023). *Machine Learning and Deep Learning: A Review of Methods and Applications*. Social Science Research Network.

<https://ssrn.com/abstract=4458723>

Sharma, N., Sharma, R., & Jindal, N. (2021). Machine Learning and Deep Learning Applications-A Vision. *Global Transitions Proceedings*, 2(1).

<https://doi.org/10.1016/j.gltcp.2021.01.004>

Soliman, G., Zaki, H., & Kilany, M. (2025). A comparative analysis of encoder only and decoder only models for challenging LLM-generated STEM MCQs using a self-evaluation approach. *Natural Language Processing Journal*, 10, 100131.

<https://www.sciencedirect.com/science/article/pii/S294971912500007X>

Sustainability Accounting Standards Board. (2018). SASB's Sustainable Industry Classification System (SICS®). <https://sasb.ifrs.org/wp-content/uploads/2018/11/SICS-Industry-List.pdf>

Sweep. (2025, juillet 2). *Sustainability data management*. Sweep.
<https://www.sweep.net/blog/what-is-double-materiality>

TCFD. (2017). *Recommendations of the Task Force on Climate-related Financial Disclosures*. <https://assets.bbhub.io/company/sites/60/2021/10/FINAL-2017-TCFD-Report.pdf>

TCFD. (2023). *Task Force on Climate-related Financial Disclosures*. TCFD.
<https://www.fsb-tcfd.org/about/>

Tombal, M. (2020, 9 avril). *Sustainability Officer chez Sibelga* [Entretien]. Bruxelles.

Toosi, A., Bottino, A. G., Saboury, B., Siegel, E., & Rahmim, A. (2021). A Brief History of AI: How to Prevent Another Winter (A Critical Review). *PET Clinics*, 16(4), 449–469. ResearchGate.

https://www.researchgate.net/publication/354387444_A_Brief_History_of_AI_How_to_Prevent_Another_Winter_A_Critical_Review

Townsend, B. (2020). From SRI to ESG: The Origins of Socially Responsible and Sustainable Investing. *The Journal of Impact and ESG Investing*, 1(1), 10–25.

<https://www.bailard.com/wp-content/uploads/2020/09/History-Socially-Responsible-Investing-and-ESG-Investing.pdf>

Union européenne. (2014, octobre 22). *Directive 2014/95/UE du Parlement européen et du Conseil du 22 octobre 2014 modifiant la directive 2013/34/UE en ce qui concerne la publication d'informations non financières et d'informations relatives à la diversité par certaines grandes entreprises et certains groupes. Journal officiel de l'Union européenne*, L 330, 1–9. <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:32014L0095>

Union européenne. (2016, avril 27). *Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données (RGPD). Journal officiel de l'Union européenne*. <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A32016R0679#d1e2188-1-1>

United Nations Global Compact. (2004). *Who Cares Wins: Connecting Financial Markets to a Changing World*.

<https://documents1.worldbank.org/curated/en/280911488968799581/pdf/113237-WP-WhoCaresWins-2004.pdf>

Vanoppen, C. (2025, 19 mars). *ESG Manager chez Lineas* [Entretien]. Bruxelles.

Wood, D. J. (1991). Corporate Social Performance Revisited. *The Academy of Management Review*, 16(4), 691–718. <https://doi.org/10.2307/258977>

Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: theory and practice. *Knowledge Engineering Review*, 10(2), 115–152.

https://www.cs.cmu.edu/~motionplanning/papers/sbp_papers/integrated1/woodridge_intelligent_agents.pdf

Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., Zhang, M., Wang, J., Jin, S., Zhou, E., Zheng, R., Fan, X., Wang, X., Xiong, L., Zhou, Y., Wang, W., Jiang, C., Zou, Y., Liu, X., & Yin, Z. (2025). *The rise and potential of large language model based agents: a survey*. arXiv. <https://arxiv.org/abs/2309.07864>

- Xu, Y., Wang, Q., An, Z., Wang, F., Zhang, L., Wu, Y., Dong, F., Qiu, C.-W., Liu, X., Qiu, J., Hua, K., Su, W., Xu, H., Han, Y., Cao, X., Liu, E., Fu, C., Yin, Z., Liu, M., & Roepman, R. (2021). *Artificial Intelligence: a powerful paradigm for scientific research. The Innovation*, 2(4). Sciencedirect.
<https://doi.org/10.1016/j.xinn.2021.100179>
- Yang, N. (2025). The Impact of GPT Models on Education: Enhancing Learning Outcomes and Addressing Challenges. *ITM Web of Conferences*, 70, 04007. https://www.itm-conferences.org/articles/itmconf/pdf/2025/01/itmconf_dai2024_04007.pdf
- Zhao, P., Cheng, N., & Jin, Z. (2023) An in-depth survey of large language model-based artificial intelligence agents. arXiv. <https://arxiv.org/pdf/2309.14365>
- Zhou, Y., Dai, S., Cao, Z., Zhang, X., & Xu, J. (2025). *Length-Induced Embedding Collapse in PLM-based Models*. arXiv . <http://arxiv.org/pdf/2410.24200>

