

RIA31 : Référentiel IA Éthique et Responsable

Institut du Numérique Responsable

(INR)

4 Octobre 2024

Avec l'intervention de :

Denis Didier
Chief Technology Officer –
Sustainable IT expert
Groupe ISIA



Fabien Abrikh
Chef de projets , Insights & Data
France at Capgemini



Laurent Bourgitteau-Guiard
Délégué Général de l'INR



Eric Benoits
Directeur de programme, Insights &
Data France at Capgemini



Et le support de :

AGENDA

- 1. Les enjeux de l'IA et du Numérique Responsable**
- 2. Présentation de la genèse de l'outil, de l'outil, des objectifs, liens ODD, méthodes de travail et contributeurs**
- 3. Démonstration et mise en oeuvre, retours d'expérience**
- 4. Next Steps**

Les enjeux de l'IA et du NR

Les enjeux de l'IA et du Numérique Responsable

Opportunités de l'IA :

- **Efficacité opérationnelle :**
 - Ciblage plus précis pour personnaliser les offres.
 - Prédiction des tendances du marché.
 - Amélioration de la prise de décision. Notamment médecine.
- **Réduction des coûts :**
 - Automatisation des tâches répétitives et chronophages (gestion des stocks, optimisation des trajets).
 - Libération des effectifs pour des activités à plus forte valeur ajoutée.
- **Expérience client améliorée :**
 - Utilisation de chatbots et assistants virtuels avancés pour une interaction plus personnalisée et interactive.

Les enjeux de l'IA et du Numérique Responsable

Risques et Précautions :

- **IA Responsable :**
 - Importance de maintenir une image de marque positive et de protéger les clients et employés
 - Nécessité d'éduquer et d'informer les utilisateurs sur les capacités et limites de l'IA
 - Nécessité de prendre en compte l'impact environnemental dans les décisions de concevoir, déployer et utiliser l'IA

➡ **Critères : PEOPLE / PLANET / PROSPERITY**

Les enjeux de l'IA et du Numérique Responsable

Catégories de Risques :

- **Hallucinations et Projections :**
 - Risque de désinformation dû à des réponses basées sur des probabilités
 - Influence des projections humaines sur les réponses de l'IA
 - Filtrage nécessaire des résultats pour éviter des conclusions erronées
- **Illusion d'Empathie :**
 - Manipulations potentielles par des réponses en première personne
 - Transfert d'états émotionnels par l'IA créant une fausse perception de performance
 - Langage naturel induisant les utilisateurs à se confier dangereusement
- **Entraînement :**
 - Risque d'utiliser des informations fausses pour entraîner l'IA
 - Diminution possible de la créativité humaine
 - Consommation énergétique importante

Les enjeux de l'IA et du Numérique Responsable

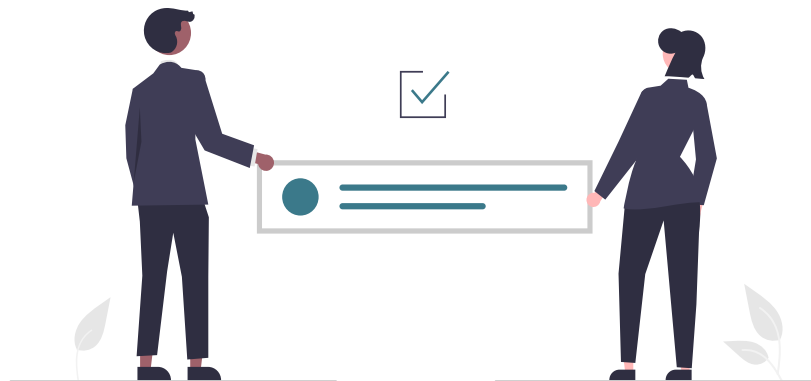
Catégories de Risques :

- **Risques Multilingues :**
 - Biais culturels dus à la surreprésentation de données en anglais
 - Nécessité de modèles multilingues équilibrés et pluriculturels
- **Problèmes Juridiques et Techniques :**
 - Absence de vérification d'âge des utilisateurs.
 - Problèmes juridiques de collecte de données
 - Difficulté d'appliquer le droit à l'oubli durant l'apprentissage.
- **Risques d'Usage :**
 - Difficulté à distinguer des productions par l'IA de celles d'un humain (filigrane - watermarks)
 - Définition de la responsabilité
 - Sur consommation énergétique

Les enjeux de l'IA et du Numérique Responsable

Conclusion :

- **Vigilance et Régulation :**
 - Collaboration nécessaire pour un avenir numérique responsable
 - Régulation essentielle pour naviguer dans le paysage technologique en évolution constante

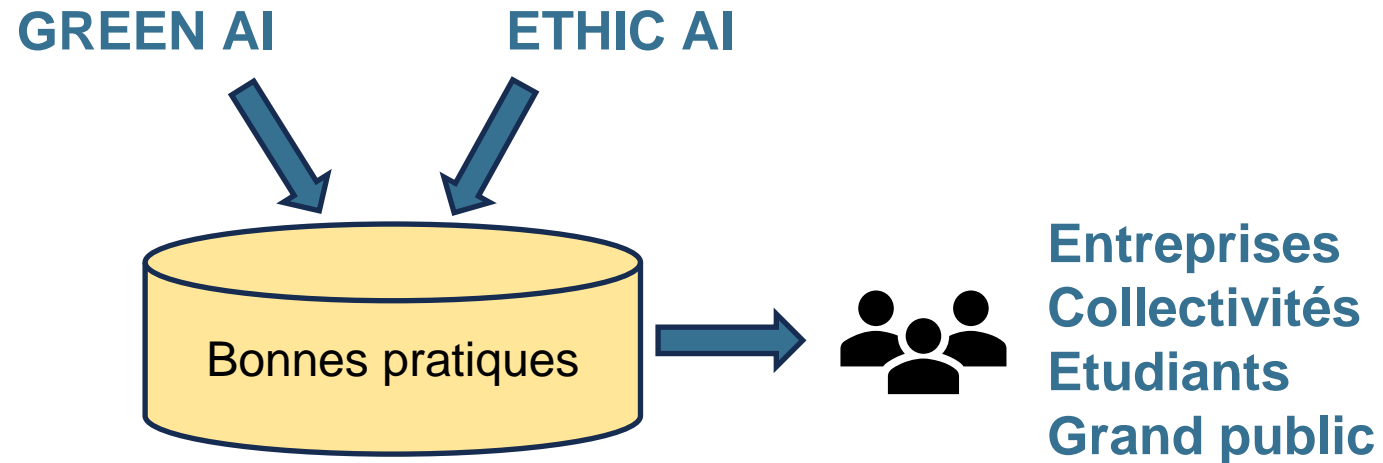


Présentation de l'outil

Génèse, objectifs, liens ODD, méthode
de travail

Objectifs

- **Centraliser les bonnes pratiques de conception, construction et exploitation**
- **Mettre à disposition au plus grand nombre d'un outil simple et compréhensible**



Bénévoles : Près de 80 pendant 2 ans

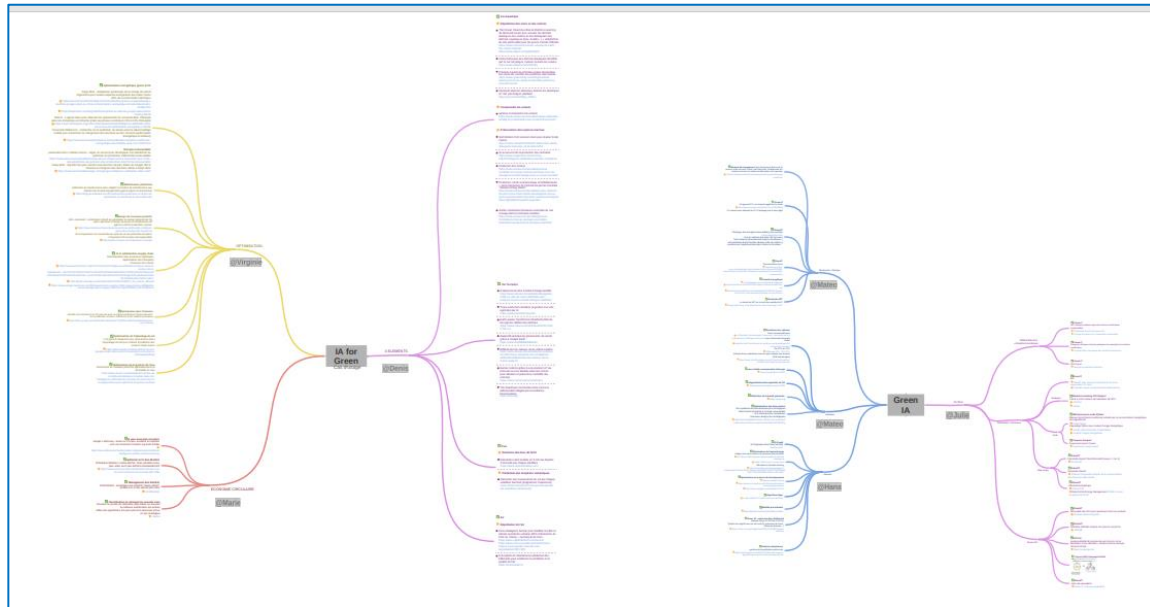
Denis	Didier
Eric	Benoits
Fabien	Abrikkh

Abdelhadi	Issoufali	Antoine	Duval	Claude	de Loupy	Francois	Auger	Herve	Leluherne
Adelaide	Kissi	Antoine	Gaget	Cleo	Tassain	Gilles	Cuyaubere	Hugo B.	Martin
Akilhoussen	Onaly	Audric	Mazzietti	Daniel	Bartolo	Guillaume	Bodiou	Ivano	Darrigo
Alberto	Tena	Aurelie	Daviaud	Denis	Debarbieux	Guillaume	Gallon	Jerome	Barre
Albina	Muzyka	Baptiste	Lauer	Doeuk	Nosing	Guillaume	Gimonnet	Julie	Trinckvel
Alexandre	Giraud	Bertrand	Lejeune	Dunand	Nicolas	Hafid Ait Sidi	Hammou	Karim	Houni
Alice	Louis	Charles	Armataffet	Duval	Antoine	Hamza	Mimi	Kathia	Couvreur
Allanic	Althenas	Christelle	Josselin	Elisabeth	Lehagre	Hana	Malha	Kevin	Le Denic
Angela	ciocan	Christophe	Porchet	Emma	Amblard	Helene	Duval	Laurène	Vaugrante
Anne	Tozzolino	Claire	Verdier	Esmâ	Talhi	Henri	Petit	Lucie	Vannieuwenhuyze
Marianne	Allanic	Nicolas	Legrand	Remy	Thibaudeau	Victor	Reutenauer		
Marie	Leduc	Nicolas	Belcourt	Richard	Bury	Vincent	De-Montalivet		
Martin	Allanic	Nicolas	Brunelet	Segolene	Dessertine-Panhard	Virginie	Boullery		
Maxime	Cornet	Nicolas	Legrand	Sergio	Garcia de Alba	Willie	Drouhet		
Maya	Guillaumont	Nicolas	Tarbes	Stephan	Mir	Yannick	Benetreau		
Mehdi	Ammi	Nicolas-N	Dunand	Sylvie	Remangeon				
Mouhamed	Gueye	Olivier	Matz	Thibaut	Duguet				
Nathali	Paz-Del-Castillo	Philippe	Broussy	Thierry	Vieville				
Nathan	Lauga	Quentin	Faulconnier	Thomas	Jean-Francois				
Nicolas	Belcourt	Raphaël	Colas	Thomas	Lefebvre				

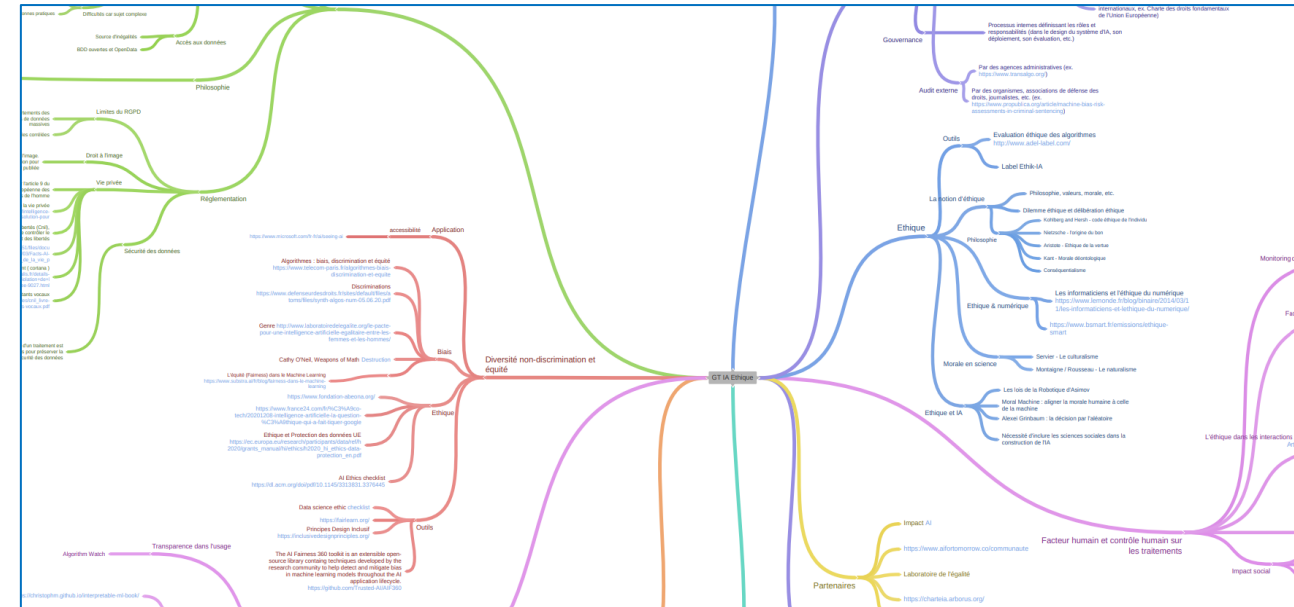
Arnaud Roudsovsky	Support Technique du site
Lisa Hurault	Chargée de communication INR

Séances d'idéation, ateliers de travail, entretiens... avec des BENEVOLES de différents horizons :
 > Data Scientist, Sociologues, Juristes, Chefs de Projets, Responsables NR ou RSE

Synthèse de l'état de l'art



AI FOR GREEN / GREEN AI



ETHIC AI

Présentation de l'outil - Grille de lecture

La structure du référentiel IA Éthique et Responsable est une arborescence qui contient :

- **5 FAMILLES**
- Des **RECOMMANDATIONS** : 13 recommandations au total
- Des **CRITERES** : plusieurs critères pour une même recommandation : 31 Critères exposés.

Présentation de l'outil

> Les Recommandations

Les RECOMMANDATIONS sont l'expression d'un objectif qui doit être adressé à cette étape projet, la formulation est présentée sous forme d'une phrase avec un verbe d'action.

Pour chaque étape projet / FAMILLE, pour certaines étapes projet, l'ensemble des thématiques peut être couvert avec un nombre de RECOMMANDATIONS réduit.

L'ordre d'apparition des RECOMMANDATIONS ne révèle pas un ordre de priorité ni d'importance.

> Les Critères

L'adhésion à la RECOMMANDATION est assortie de CRITÈRES qui font l'objet d'une évaluation exprimée au travers d'une question.

Pour une RECOMMANDATION donnée, l'ordre d'apparition des CRITÈRES n'introduit aucune notion de priorité.

> Les catégories de CRITÈRES

Les CRITÈRES sont présents pour questionner une RECOMMANDATION sous plusieurs "facettes". Certaines de ces questions appelle une réponse formelle et une mise en œuvre concrète

Présentation de l'outil

> Liaisons avec les Objectifs de Développement Durable de l'ONU (ODD)

Les sujets couverts par la RECOMMANDATION ou LE CRITÈRE peuvent être rattachés à un ou plusieurs des 17 ODD.

Pour chaque RECOMMANDATION/CRITÈRE si l'un des ODD est adressé, bien sûr très partiellement, l'ODD concerné est exposé.

Pour certaines organisations qui ont relié leur stratégie RSE/CSR aux ODD de l'ONU, cette mise en relation des enjeux NR et des ODD pris en compte par la RSE/CSR peut être utile.

> Impact PLANET / PEOPLE / PROSPERITY

Afin de positionner les bénéfices de cet élément par rapport aux 3 dimensions du Numérique Responsable : Planet, People, Prosperity, l'évaluation de l'impact est présentée par « x » dans la colonne correspondante

Présentation de l'outil

> Difficulté

Chaque critère est accompagné d'une évaluation de la difficulté de mise en œuvre : simple, moyen, difficile.

> Récurrence

Cette information répond à la question "A quel moment prendre en compte cet élément ?"

- Certain CRITÈRES peuvent n'être adressé qu'une seule fois, le terme "Once" est utilisé.
- Certain CRITÈRES doivent être systématiquement pris en compte à chaque mise à jour de fonctionnalité du service : "Update" identifie ces cas.
- Lorsque le contexte extérieur peut modifier les éléments pris en compte par le CRITERE, il est nécessaire de vérifier périodiquement ce CRITÈRE, le terme "Periodic" est utilisé. La fréquence n'est pas précisée mais laissée au libre arbitre et au bon sens.

> Tests

Des tests sont ensuite proposés pour préciser les axes d'investigation à couvrir.

Présentation de l'outil

> Argumentaire

Pour affirmer la nécessité de ces informations dans le référentiel, une information de précision est indiquée. Cet élément peut aider à comprendre pourquoi ce point est présent et son importance dans la démarche.

> Use Case / Limites

Afin de vérifier le niveau de conformité avec ce CRITÈRE les éléments qui peuvent être utilisés sont présentés dans ce bloc ou les limites identifiées.

> Les métiers et acteurs

Le CRITÈRE est plutôt destiné à un public ciblé pour qu'il soit pris en compte.
Cette classification permet de ventiler les CRITÈRES vers les acteurs les plus à même de le traiter.
Les acteurs sont l'un ou plusieurs des profils suivants :

- Développeurs
- Designers
- Gestion de projet
- Product Owner / métier
- Encadrement / manager
- RSE / NR
- Ressources Humaines
- DSI
- Toute l'équipe projet
- Architectes d'entreprise
- Data Scientists

Démonstration et mise en oeuvre

RIA31, le Référentiel IA Ethique et Responsable

5 Familles | 13 recommandations – 31 critères

RIA31 Guide IA Ethique et Responsable



STRATEGIE

L'étape de stratégie projet permet de déterminer la pertinence et les enjeux du projet.

CADRAGE PROJET

Regroupent les éléments de cadrage projet, les moyens mis en oeuvre, les objectifs et contraintes du projet sur toute la durée de vie du produit cible. Indépendamment du type de gestion projet : AGILE ou cycle en V classique.

EXPERIENCE UTILISATEUR

Les étapes et méthodes de conception des services numériques pour définir les meilleures solutions d'interactions avec l'utilisateur.

INTERFACE UTILISATEUR

Ensemble des composants en opération sur un terminal utilisateur pour permettre l'utilisation d'un service IA.

SOCLE TECHNOLOGIQUE

Le socle technologique représente la traduction informatique des processus métiers, les moyens techniques et données mis en oeuvre pour leur utilisation, ainsi que toutes les interactions externes mises en oeuvre pour leur réalisation.

Retours d'expérience de nos bêta testeurs

Nos bêta testeurs

5 organisations membres de l'INR ont été nos beta testeurs durant 1,5 mois sur des projet IA existants ou en construction.

Open.
Studio

sopra  steria

 **BNP PARIBAS**
PERSONAL FINANCE 


norsys

Capgemini 

Les retours d'expérience

- Adresse les enjeux de transparence pour les modèles de langage perçus comme des "boîtes noires".
- Critères clairs qui sensibilisent aux pratiques durables et éthiques
- Permet de prendre du recul et d'identifier les éléments manquants pour renforcer les valeurs de numérique responsable



REX OPENSTUDIO

Que vous a apporté ce référentiel et quels conseils ou trucs et astuces souhaitez-vous partager à la communauté pour un usage efficient du référentiel ?

- Nous a permis de nous **confronter à un regard extérieur pour améliorer nos processus** vers une IA plus responsable et éthique.
- **L'appel à une personne extérieure au projet afin de l'évaluer.**
- **Une meilleure compréhension de chaque critère et argumentaire avant de tester et évaluer les indicateurs.**



REX SOPRASTERIA

Que vous a apporté ce référentiel et quels conseils ou trucs et astuces souhaitez-vous partager à la communauté pour un usage efficient du référentiel ?

- **Bonne réception par l'équipe** et qui permet de se sensibiliser sur les pratiques d'une IA durable et éthique.
- **Se poser les bonnes questions** pour évaluer l'impact et la conformité éthique d'un service numérique utilisant l'IA en moins de 3h.
- **Ne pas chercher à appliquer tous les critères.** Utiliser la notion d'impact 3P et la notion de difficulté de mise en œuvre pour prioriser.
- Distinguer ce qui est applicable à **l'IA classique vs IA générative.**
- Réaliser l'évaluation avec une **équipe pluridisciplinaire.**
- **Complexité de l'évaluation de la partie éthique** sur les modèles de langages qui sont devenus des boîtes noires → Enjeu de transparence des modèles



Que vous a apporté ce référentiel et quels conseils ou trucs et astuces souhaitez-vous partager à la communauté pour un usage efficient du référentiel ?

- **Pour s'appuyer sur un référentiel travaillé collectivement est une force** (ex: framework écoconception de l'INR).
- Conseil (que nous mettrons certainement en pratique): **s'approprier le référentiel, étudier comment l'incorporer dans la gouvernance existante de l'entreprise et le vendre comme un atout et non une contrainte**

REX NORISYS



Les apports de cette démarche de Beta Test

- **Prise de conscience du large spectre** des domaines en lien avec l'Intelligence Artificielle
- **Du recul sur l'impact de ce que nous mettons en place**
- Des actions pour nous améliorer (sécurité, impact environnemental, transparence, ...)

Le référentiel INR :

- **Une structure rassurante avec des questions accessibles (ni trop technique, ni trop simple)**
- **En cohérence avec les autres référentiels de l'INR**

Un conseil : Faire intervenir « un candide » pour challenger ses propres justifications



REX CAPGEMINI

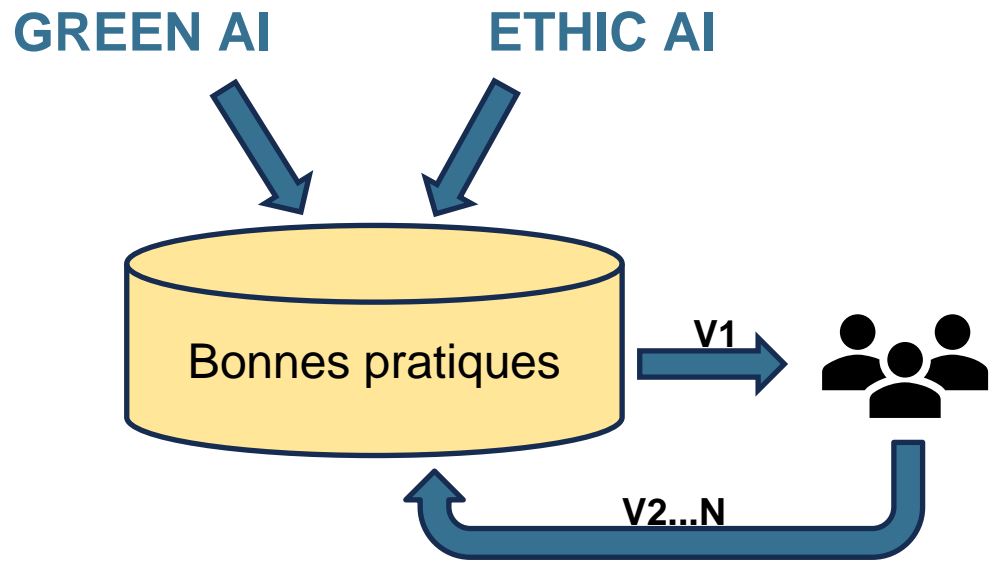


Que vous a apporté ce référentiel et quels conseils ou trucs et astuces souhaitez-vous partager à la communauté pour un usage efficient du référentiel ?

Le conseil principal que nous retenons est **d'avoir connaissance des critères dès les premières phases du projet**, même sans chercher à les appliquer. Ensuite, il nous semble important de prendre du recul, certains critères ne sont pas applicables à tous les projets.



Principe de “Work In Progress” permanent



Positionnement identique au GR491.

V1 : Approbation de la V1 par les beta testeurs.

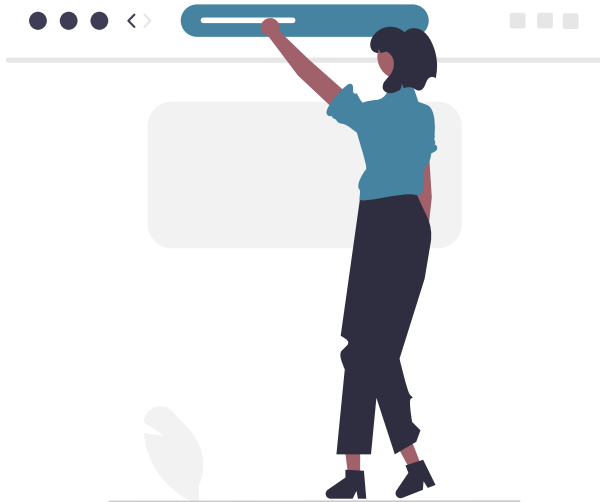
V2...N : Nécessité de faire vivre le référentiel :

- Avancées technologiques
- Maturité des acteurs
- Réglementations
- Précisions / ajustements

Toutes vos suggestions d'amélioration, de nouvelles recommandations, de nouveaux critères :

 gtia@institutnr.org

Le RIA31



Découvrez le RIA31

<https://ref-ia.isit-europe.org/>

 Suivez nous sur LinkedIn : [Institut du Numérique Responsable](#)

 Contribuez au RIA31 : gtia@institutnr.org