LIVRE BLANC

Du dirigeant au double numérique

Capture, réalisme, fiabilité et consentement des humains numériques MetaHuman en contexte d'entreprise

Revue de la littérature, cadre juridique et méthode

Thomas Varin

Directeur technique

VP Capture Studio

Montréal (Québec), Canada

info@sosie3d.com · sosie3d.com · kiosk3d.com

Version 1.4 — 17 novembre 2026

Citation suggérée : Varin, T. (2026). *Du dirigeant au double numérique : capture, réalisme, fiabilité et consentement des humains numériques MetaHuman en contexte d'entreprise* [Livre blanc]. VP Capture Studio.

Avant-propos

En développant notre technologie de capture à Montréal, nous nous sommes heurtés aux trois mêmes questions que se pose toute organisation qui envisage de confier son image à un double numérique. Est-ce que le résultat paraîtra étrange ? Est-ce que le personnage peut dire n'importe quoi ? Et a-t-on le droit d'utiliser le visage et la voix d'une personne ?

J'ai voulu y répondre avec des sources plutôt qu'avec des promesses. Ce livre blanc rassemble ce que la recherche, le droit et la documentation technique disent aujourd'hui de ces trois questions, et la façon dont nous en avons tiré notre méthode de travail. Là où les données manquent, nous le disons ; la section 10 décrit ce que nous mesurerons nous-mêmes.

Il s'adresse aux dirigeants, aux équipes de communication et aux juristes qui envisagent un double numérique. J'espère qu'il les aidera à poser les bonnes questions, à nous comme à tout autre fournisseur.

Vos commentaires et vos critiques sont les bienvenus à info@sosie3d.com : ils nourriront la prochaine version.

Thomas Varin
Directeur technique, VP Capture Studio
Montréal, le 17 novembre 2026

Résumé

Les humains numériques photoréalistes sont passés en peu de temps du laboratoire aux usages d'entreprise : des dirigeants présentent déjà leurs résultats financiers par avatar, et les outils de création comme MetaHuman sont désormais ouverts à l'usage commercial et en partie publiés en code source libre. Ce livre blanc examine trois questions : (1) le réalisme visuel d'un double numérique favorise-t-il la confiance du public ? (2) à quelles conditions un double conversationnel est-il fiable et perçu comme tel ? (3) quel cadre juridique et éthique encadre la reproduction du visage et de la voix d'une personne réelle ?

Une revue narrative de sources primaires publiées entre 2022 et novembre 2026 — études révisées par des pairs, textes juridiques, décisions et documentation technique officielle — conduit à trois constats. Premièrement, le réalisme élevé est associé à une confiance plus grande, alors que le réalisme intermédiaire est le plus dérangeant. Deuxièmement, la fiabilité perçue d'un agent conversationnel dépend surtout de sa capacité à s'abstenir lorsque ses sources sont insuffisantes et à masquer la latence par des comportements naturels. Troisièmement, malgré la diversité des régimes québécois, européen et américain, le consentement écrit, précis, distinct et révocable s'impose comme norme commune. Nous en tirons une méthode de production en sept étapes, dix principes de bonne pratique, une grille d'évaluation des fournisseurs et un programme de mesure.

Mots-clés : *humain numérique ; double numérique ; MetaHuman ; photogrammétrie ; agent conversationnel incarné ; voix clonée ; confiance ; vallée dérangeante ; Loi 25 ; consentement ; provenance des contenus*

Avis

Les constats factuels de ce document renvoient à des sources publiques vérifiées, citées selon les normes de l'APA (7e édition). Les estimations de marché sont reproduites telles que publiées par leurs auteurs, sans être endossées. L'état du droit est décrit au 17 novembre 2026 ; ce document ne constitue pas un avis juridique. MetaHuman et Unreal Engine sont des marques de commerce d'Epic Games, Inc. ; VP Capture Studio utilise ces outils sous licence et n'est pas affilié à Epic Games.

Table des matières

1. Introduction 6

1.1 Contexte 6

1.2 Questions de recherche 6

1.3 Contribution et plan 7

2. Définitions et périmètre 7

3. Méthode de la revue 8

4. État de la technologie 8

4.1 Capture du visage et de la tête 8

4.2 Du modèle numérisé au personnage animable 9

4.3 Animation du visage, du corps et de la voix 9

4.4 Rendu 10

4.5 Chaîne conversationnelle 10

5. Revue de la littérature : perception, voix et conversation 10

5.1 La vallée dérangeante, de l'hypothèse aux données (QR1) 10

5.2 Photoréalisme, confiance et détection 11

5.3 La voix clonée 11

5.4 Présence sociale 11

5.5 Agents incarnés en milieu réel (QR2) 11

5.6 Latence et remplissages naturels 12

5.7 La vidéo personnalisée 12

5.8 Synthèse 12

6. Cadre juridique et éthique (QR3) 13

6.1 Québec : vie privée, image et voix 13

6.2 Renseignements personnels et biométrie 13

6.3 Responsabilité pour les propos de l'agent 13

6.4 Canada 14

6.5 International 14

6.6 Normes professionnelles 14

6.7 Provenance et prévention de la fraude 15

6.8 Enjeux éthiques 15

7. Marché et usages en entreprise 15

7.1 Estimations publiées 15

7.2 Signaux du marché 16

7.3 Familles d'usage 16

8. Méthode VP Capture 17

8.1 Cadrage et consentements 17

8.2 Capture 17

8.3 Reconstruction, conformation et validation....17
8.4 Fiabilité du double conversationnel....18
8.5 Déploiement, transparence et fin de vie....19
9. Discussion....19
9.1 Réponses aux questions de recherche....19
9.2 Dix principes de bonne pratique....19
9.3 Grille d'évaluation d'un fournisseur....20
10. Limites et programme de mesure....20
11. Conclusion....21
Glossaire....21
Références....22

1. Introduction

1.1 Contexte

En mai 2025, deux entreprises cotées ont confié à l'avatar de leur président la présentation de leurs résultats trimestriels. Celui de Klarna a ouvert une vidéo de 83 secondes en se présentant comme l'avatar IA du président ; celui de Zoom a livré les remarques d'ouverture avant que le vrai dirigeant ne réponde aux questions en direct (Business Today, 2025 ; TechCrunch, 2025). L'accueil a été partagé : des observateurs ont relevé un regard figé et un clignement mécanique (TechSpot, 2025). Ces deux épisodes résument l'enjeu : l'usage existe, mais sa crédibilité dépend de la qualité d'exécution.

La même technologie sert aussi à la fraude. En janvier 2024, un employé du bureau de Hong Kong de la firme d'ingénierie Arup a effectué 15 virements totalisant 200 millions de dollars de Hong Kong après une vidéoconférence dans laquelle le chef des finances et plusieurs collègues étaient des imitations numériques construites à partir de vidéos publiques (CFO Dive, 2024 ; Dezeen, 2024). Un double numérique légitime doit donc se distinguer nettement d'un hypertrucage : par le consentement de la personne représentée, par la transparence envers le public et par la traçabilité des contenus.

Entre ces deux réalités, les outils ont changé d'échelle. MetaHuman, l'environnement d'Epic Games pour créer et animer des humains numériques, est sorti de l'accès anticipé en 2025 avec une licence qui en permet l'usage commercial hors d'Unreal Engine (Thacker, 2025), et sa version 5.8 publie en code source libre une partie de ses bibliothèques fondamentales (Epic Games, 2026c). Ce qui exigeait un studio d'effets visuels devient accessible à des organisations de taille moyenne. La question n'est plus de savoir si l'on peut produire un double crédible, mais comment le faire de manière fiable, licite et digne de confiance.

1.2 Questions de recherche

Ce livre blanc poursuit trois questions :

1. **QR1.** Le réalisme visuel d'un double numérique favorise-t-il ou compromet-il la confiance du public ?
2. **QR2.** À quelles conditions un double conversationnel, qui répond à voix haute, est-il fiable et perçu comme tel ?
3. **QR3.** Quel cadre juridique et éthique encadre la reproduction du visage et de la voix d'une personne réelle, au Québec et chez ses principaux partenaires ?

1.3 Contribution et plan

Notre contribution est double : une synthèse à jour des connaissances techniques, scientifiques et juridiques ; et une méthode opérationnelle qui en découle, appliquée dans notre studio de Montréal. La section 2 précise les définitions ; la section 3, la méthode de revue ; la section 4 décrit l'état de la technologie ; la section 5 passe en revue la littérature sur la perception et la conversation ; la section 6 analyse le cadre juridique et éthique ; la section 7 examine le marché et les usages ; la section 8 présente notre méthode ; la section 9 discute les résultats et propose des principes de bonne pratique ; la section 10 expose les limites et le programme de mesure.

2. Définitions et périmètre

Le vocabulaire du domaine est instable : « avatar », « humain virtuel », « humain numérique », « double », « réplique » et « hypertrucage » désignent des objets voisins mais distincts. Les régimes juridiques s'appuyant sur des définitions précises, nous retenons celles du tableau 1.[1]

Tableau 1

Définitions retenues

Terme	Définition	Source
Humain numérique	Personnage de synthèse photoréaliste, animé et, le cas échéant, conversationnel.	Usage courant
Double numérique	Humain numérique qui reproduit l'apparence, le mouvement et, le cas échéant, la voix d'une personne réelle, avec son consentement.	Le présent document
Réplique numérique	Représentation électronique très réaliste, générée par ordinateur, facilement identifiable comme la voix ou l'apparence d'une personne.	Californie, AB 2602 et AB 1836
Hypertrucage (deep fake)	Contenu image, audio ou vidéo généré ou manipulé par IA, qui ressemble à des personnes, objets, lieux, entités ou événements existants et paraîtrait faussement authentique.	UE, AI Act, art. 3(60)
Agent conversationnel incarné	Agent logiciel doté d'un corps visible, qui dialogue avec la parole, le regard et les gestes.	Cassell et al. (2000)

Note. Les définitions de l'AI Act et des lois californiennes sont résumées ; voir Zolynski (2026) et Davis+Gilbert (2024) pour le texte intégral.

La distinction centrale est celle du consentement : un double numérique et un hypertrucage peuvent être techniquement identiques ; seul le consentement de la personne, et la transparence envers le public, les séparent. Notre périmètre couvre les doubles de personnes réelles en contexte

[1] Nous employons « double numérique » pour la représentation d'une personne réelle, consentante, et « humain numérique » pour toute représentation photoréaliste, y compris de personnages inventés.

d'entreprise — dirigeants, porte-parole, formateurs — et exclut les personnages entièrement inventés, les personnes décédées et les usages politiques.

3. Méthode de la revue

Nous avons réalisé une revue narrative ciblée en novembre 2026, portant principalement sur la période 2022–2026 et complétée par des travaux fondateurs plus anciens lorsqu'ils structurent le domaine (Mori et al., 2012 ; Reeves et Nass, 1996). Trois familles de sources ont été retenues : (a) des études empiriques révisées par des pairs ou présentées en conférence avec comité de lecture (PNAS, Psychological Science, Scientific Reports, Frontiers in Psychology, Journal of Science Communication, ICLR, CVPR, ACM CUI, ISPRS) ; (b) des textes législatifs, décisions et documents d'autorités publiques ; (c) la documentation technique officielle des éditeurs. La presse spécialisée n'a servi qu'à dater des événements et a été systématiquement recoupée.

Deux critères d'exclusion ont été appliqués : tout chiffre dont la source primaire n'a pu être retrouvée, et toute affirmation juridique dont le stade législatif n'a pu être vérifié. Ce second critère s'est révélé nécessaire : plusieurs publications présentent comme adoptés des textes qui ne sont encore que des projets (section 6.5). Il s'agit d'une revue narrative et non systématique ; ses limites sont discutées à la section 10.

La recherche documentaire et la rédaction ont été réalisées avec l'aide d'outils d'intelligence artificielle générative. L'auteur a vérifié les sources citées et assume l'entière responsabilité du contenu.

4. État de la technologie

La production d'un double numérique enchaîne quatre opérations : la capture de la personne, la construction d'un personnage animable, son animation (visage, corps, voix) et son rendu. Lorsqu'il est conversationnel, s'y ajoute une chaîne de dialogue. Nous examinons chaque maillon.

4.1 Capture du visage et de la tête

La photogrammétrie reconstruit une surface 3D à partir de nombreuses photographies prises sous des angles différents : la structure acquise par le mouvement (structure-from-motion) estime la position des caméras, puis la stéréo multivue densifie la géométrie (Schönberger et Frahm, 2016). Pour un visage humain, la capture doit être simultanée, car le moindre mouvement entre deux prises dégrade la reconstruction ; d'où l'usage de réseaux de caméras déclenchées ensemble, souvent avec un flash pour figer l'expression.

Ces réseaux ne requièrent plus de matériel de cinéma. Giuliani et al. (2024) ont évalué un système multivue fondé sur des caméras à capteur Sony IMX477 (12,3 mégapixels) et objectifs de 16 mm pour

produire des modèles 3D du visage : avec 27 points de vue répartis sur trois rangées et dix mesures d'échelle, l'écart-type atteint 0,9 mm et 80 % des points se situent à ±1 mm d'un modèle de référence.[2] Les auteurs présentent toutefois leur travail comme une étude de faisabilité : la précision sous le millimètre ne vaut que pour 80 % des points, et les essais sur volontaires humains n'ont été évalués que qualitativement. Kersten et al. (2024) décrivent pour leur part un système de 24 caméras du même type, à déclenchement synchrone. Ces travaux établissent qu'une capture multivue rigoureuse, avec un calibrage d'échelle soigné, approche le millimètre à coût maîtrisé.

4.2 Du modèle numérisé au personnage animable

Un nuage de points n'est pas un personnage. Pour parler et exprimer des émotions, le modèle doit être « conformé » à une topologie standard munie d'un rig — un squelette et un ensemble de commandes faciales. Les rigs faciaux modernes héritent du système de codage des actions faciales (FACS), qui décompose les expressions en unités d'action musculaires (Ekman et Friesen, 1978). MetaHuman fournit un tel rig standardisé : la fonction Mesh to MetaHuman conforme un modèle numérisé à la topologie MetaHuman, et la version 5.8 étend cette conformation au corps entier, ce qui convient précisément aux personnages issus de numérisations (Epic Games, 2026b).

Le cadre de licence a évolué en parallèle. Depuis la version 5.6 (juin 2025), l'usage de MetaHuman relève de la licence d'Unreal Engine ; les personnages peuvent être utilisés dans d'autres moteurs et logiciels, y compris dans des projets commerciaux, et seuls les studios dont les revenus annuels dépassent un million de dollars américains doivent acquérir des licences par poste (Epic Games, s.d. ; Thacker, 2025). L'application web MetaHuman Creator a été abandonnée en mars 2026, la création étant désormais intégrée à Unreal Engine (Epic Games, 2026a). Enfin, Epic a publié sous licence MIT les bibliothèques RigLogic et DNA, qui évaluent le rig facial MetaHuman en dehors d'Unreal Engine (Epic Games, 2026c) — un gage de pérennité pour un personnage dont la vie utile dépasse celle d'une version de logiciel. Unreal Engine 5.8 est d'ailleurs présenté comme la dernière version majeure prévue d'UE5 avant UE6 (Epic Games, 2026d).

4.3 Animation du visage, du corps et de la voix

Trois sources d'animation coexistent. La première est la performance captée : MetaHuman Animator peut désormais capter la performance complète — visage, corps ou les deux — à partir d'une seule caméra, et l'animation faciale en temps réel accepte un flux vidéo en continu (Epic Games, 2026b). La deuxième est la capture du mouvement sans marqueurs par plusieurs caméras : des logiciels libres comme Pose2Sim triangulent des détections de pose 2D pour obtenir une cinématique 3D du corps entier (Pagnon et al., 2022). La troisième est l'animation pilotée par la voix : en septembre 2025, NVIDIA a publié en code source libre les modèles, la trousse de développement et le cadre

[2] Dans l'étude, trois caméras seulement ont été déplacées sur neuf positions pour simuler les 27 points de vue ; un système réel à déclenchement simultané élimine en outre le mouvement du sujet entre les prises.

d'entraînement d'Audio2Face, qui produit l'animation faciale et la synchronisation labiale à partir de l'audio (Animation World Network, 2025).

La voix elle-même peut être clonée à partir d'enregistrements de la personne. Comme nous le verrons (section 5.3), le public distingue mal une voix clonée de la voix réelle ; la voix doit donc être traitée comme un attribut aussi identifiant que le visage, tant pour la sécurité des données que pour le consentement.

4.4 Rendu

Le rendu en temps réel, au cœur d'Unreal Engine, permet l'interaction mais impose un budget de calcul par image ; le rendu précalculé atteint une qualité supérieure mais exclut l'improvisation. Une troisième voie émerge : le « Gaussian splatting » représente une scène par des millions de gaussiennes 3D rendues en temps réel (Kerbl et al., 2023), et ses variantes produisent des têtes photoréalistes ré-éclairables, y compris les cheveux et les reflets fins (Saito et al., 2024 ; Zhang et al., 2025). À court terme, ces approches complètent le personnage riggé plutôt qu'elles ne le remplacent : le rig standardisé reste l'atout pour l'animation dirigée et l'interopérabilité.

4.5 Chaîne conversationnelle

Un double conversationnel enchaîne la reconnaissance de la parole, la recherche d'information, la rédaction de la réponse, la synthèse vocale et l'animation. La génération augmentée par recherche (retrieval-augmented generation) consiste à fournir au modèle de langage des passages extraits d'une base documentaire au moment de répondre, plutôt que de s'appuyer sur ses seuls paramètres (Lewis et al., 2020). Cette architecture réduit l'invention de faits, sans l'éliminer : lorsque le contexte récupéré est insuffisant, les modèles ont tendance à produire une réponse plutôt qu'à s'abstenir (Joren et al., 2025). Chaque maillon ajoute en outre de la latence, dont nous verrons l'effet sur la perception (section 5.6). La figure 2 (section 8) présente l'architecture que nous en déduisons.

5. Revue de la littérature : perception, voix et conversation

5.1 La vallée dérangeante, de l'hypothèse aux données (QR1)

Formulée par le roboticien Masahiro Mori en 1970, l'hypothèse de la « vallée dérangeante » prédit que l'affinité envers une figure humanoïde croît avec sa ressemblance humaine, puis chute brusquement lorsque la figure est presque, mais pas tout à fait, humaine (Mori et al., 2012). L'hypothèse a longtemps servi d'argument contre le photoréalisme. Les données récentes la nuancent fortement.

Une méta-analyse en réseau, préenregistrée, portant sur 13 études publiées entre 2015 et 2025 et totalisant 2 343 participants, conclut que les avatars au réalisme élevé sont jugés plus attirants et plus dignes de confiance que les avatars au réalisme moyen ou faible, et que c'est le réalisme moyen qui

suscite le plus de malaise (Tao et al., 2025).[3] Une expérience menée auprès de 491 personnes, dans un contexte de communication scientifique, va dans le même sens : plus l'avatar est réaliste, plus il est jugé digne de confiance, sans effet de vallée dérangeante ; la familiarité avec l'IA et la confiance envers l'institution modulent cette perception (Baake et al., 2025).

5.2 Photoréalisme, confiance et détection

Pour les images fixes, la frontière est déjà franchie. Des visages générés par réseaux antagonistes se sont révélés indiscernables de vrais visages et ont été jugés en moyenne 7,7 % plus dignes de confiance (Nightingale et Farid, 2022). Miller et al. (2023) observent même un « hyperréalisme » — des visages de synthèse jugés plus humains que de vrais visages — mais seulement pour les visages blancs, ce qu'ils attribuent à la surreprésentation de ces visages dans les données d'entraînement ; les participants les plus trompés étaient aussi les plus confiants. Deux leçons en découlent : la perception de l'authenticité n'est plus un garde-fou fiable, et la qualité d'un rendu peut varier selon les caractéristiques de la personne représentée.

5.3 La voix clonée

Barrington et al. (2025) montrent que des voix clonées à partir de courts enregistrements, avec un outil commercial grand public, sont perçues comme ayant la même identité que la voix réelle environ 80 % du temps, et ne sont reconnues comme artificielles qu'environ 60 % du temps ; les performances s'améliorent pour des extraits plus longs. La voix est donc un attribut d'identité à part entière, qui appelle le même niveau de consentement et de protection que le visage.

5.4 Présence sociale

La théorie de l'équation médiatique soutient que les personnes appliquent spontanément aux médias et aux ordinateurs les règles sociales qu'elles appliquent aux humains (Reeves et Nass, 1996). Un personnage qui regarde, parle et gesticule déclenche ces règles plus fortement qu'un texte. C'est l'atout du double numérique — une présence que le texte ne procure pas — et sa responsabilité : une erreur prononcée par un visage familier engage davantage qu'une erreur affichée à l'écran.

5.5 Agents incarnés en milieu réel (QR2)

Les agents conversationnels incarnés sont étudiés en milieu réel depuis plus de quinze ans, notamment comme guides de musée : les jumelles virtuelles Ada et Grace accueillent les visiteurs du Museum of Science de Boston depuis 2009 (Traum et al., 2012). Une évaluation récente d'un avatar conversationnel déployé dans un musée scientifique, auprès de 118 visiteurs, rapporte que les visiteurs interrogent souvent l'avatar sur lui-même et que ses réponses sont de bonne qualité environ

[3] La méta-analyse classe les niveaux de réalisme à l'aide de la surface sous la courbe de classement cumulatif (SUCRA) ; le protocole a été préenregistré sur l'Open Science Framework.

la moitié du temps ; les auteurs soulignent l'importance de maintenir la technologie sous-jacente à un haut niveau (InformalScience, 2025). Ce taux rappelle que l'engagement ne vaut rien sans la justesse.

5.6 Latence et remplissages naturels

Dans une étude intrasujets de 54 participants interagissant en réalité virtuelle avec des agents propulsés par un grand modèle de langage, dans des scénarios de magasin, d'hôtel et de musée, Maslych et al. (2025) ont comparé des délais de 1,5, 4 et 6,5 secondes. Le délai dégrade la plupart des mesures, et les auteurs concluent qu'une latence de plus de 4 secondes détériore l'expérience. Les « remplissages naturels » — un geste de réflexion accompagné d'une courte phrase — améliorent nettement le temps de réponse perçu aux délais moyens et longs, alors qu'un indicateur artificiel (icône et son de chargement) n'apporte aucune amélioration significative. Les participants se disent plus disposés à utiliser un système lent lorsque des remplissages sont présents, tout en préférant les réponses rapides.

5.7 La vidéo personnalisée

Dans un essai randomisé sur le terrain auprès de 21 328 clients d'un détaillant, des publicités vidéo personnalisées produites par IA générative ont obtenu un taux d'engagement de 6 à 9 points de pourcentage supérieur à celui d'images personnalisées et de vidéos génériques, avec d'importantes économies de production (Kapoor et Kumar, 2025). Dennis et Kizer (2025) observent aussi davantage de téléchargements, de revisionnements et de clics pour les vidéos personnalisées. Ces résultats portent sur le marketing grand public et sur une exposition unique ; leur transposition aux communications internes — les vœux d'un président à chaque employé, par exemple — reste à mesurer.

5.8 Synthèse

Tableau 2

Synthèse des constats et implications pour la conception

Question	Constat principal	Implication
Réalisme (QR1)	Réalisme élevé = confiance accrue ; réalisme moyen = malaise maximal (Tao et al., 2025 ; Baake et al., 2025).	Viser le réalisme élevé ou s'abstenir.
Détection	Le public distingue mal le vrai du synthétique, visages comme voix (Nightingale et Farid, 2022 ; Barrington et al., 2025).	La transparence incombe au producteur.
Fiabilité (QR2)	Sans contexte suffisant, les modèles inventent plutôt que s'abstenir (Joren et al., 2025).	Abstention explicite et sources approuvées.
Latence (QR2)	Au-delà de 4 s, l'expérience se dégrade ; les remplissages naturels aident (Maslych et al., 2025).	Budget < 4 s et attente « habitée ».
Personnalisation	+6 à 9 points d'engagement en marketing (Kapoor et Kumar, 2025).	Prometteur ; à mesurer en interne.

6. Cadre juridique et éthique (QR3)

6.1 Québec : vie privée, image et voix

Au Québec, le droit au respect de la vie privée est protégé par la Charte des droits et libertés de la personne (art. 5) et par le Code civil, qui considère notamment comme une atteinte le fait d'utiliser le nom, l'image, la ressemblance ou la voix d'une personne à toute autre fin que l'information légitime du public (C.c.Q., art. 35-36). Un double numérique utilisé à des fins commerciales requiert donc, en principe, le consentement de la personne représentée.

6.2 Renseignements personnels et biométrie

La Loi sur la protection des renseignements personnels dans le secteur privé, modernisée par la Loi 25, exige que le consentement soit manifeste, libre, éclairé et donné à des fins spécifiques, et qu'il soit demandé à chacune de ces fins (art. 14 ; Usercentrics, 2026). La Loi concernant le cadre juridique des technologies de l'information exige en outre le consentement exprès pour vérifier ou confirmer l'identité d'une personne par biométrie, et la déclaration à la Commission d'accès à l'information de toute banque de caractéristiques biométriques au plus tard 60 jours avant sa mise en service (LCCJTI, art. 44-45 ; Gowling WLG, s.d.). La Commission applique ces règles avec rigueur, comme l'a montré sa décision sur un projet de reconnaissance faciale d'un détaillant (Torys, 2025).

Un double numérique ne sert pas à identifier une personne ; il est néanmoins construit à partir de mesures précises de son visage. Par prudence, nous traitons ces données comme des renseignements sensibles, et l'application des articles 44 et 45 à chaque projet doit être validée par un conseiller juridique.[4]

6.3 Responsabilité pour les propos de l'agent

Dans Moffatt c. Air Canada (2024), le transporteur a été tenu responsable d'une information erronée sur les tarifs de deuil fournie par l'agent conversationnel de son site ; le tribunal a rejeté l'argument voulant que l'agent soit une entité distincte, responsable de ses propres actes, et a conclu à une déclaration inexacte négligente.[5] Pour un double numérique, la conséquence est directe : l'entreprise répond de ce que dit le visage de son président ou de son porte-parole comme de toute autre communication officielle.

[4] La décision Moffatt émane du Tribunal de résolution civile de la Colombie-Britannique (petites créances) ; elle n'a pas valeur de précédent obligatoire au Québec, mais illustre le raisonnement applicable à la déclaration inexacte négligente.

[5] L'application des articles 44 et 45 de la LCCJTI dépend de la finalité : ils visent la vérification ou la confirmation de l'identité et les banques de caractéristiques biométriques. La qualification de chaque projet relève d'un avis juridique.

6.4 Canada

Au fédéral, le projet de Loi sur l'intelligence artificielle et les données, partie du projet de loi C-27, est mort au Feuilleton à la prorogation du Parlement le 6 janvier 2025 ; aucun texte ne l'a remplacé, et le Canada n'a toujours pas de loi fédérale générale sur l'IA (CASRAI, 2026). L'encadrement repose donc sur le droit existant de la vie privée, de la responsabilité civile et de la protection du consommateur.

6.5 International

Dans l'Union européenne, les obligations de transparence de l'article 50 de l'AI Act s'appliquent depuis le 2 août 2026 : les déployeurs doivent divulguer qu'un hypertrucage a été généré ou manipulé artificiellement, et les fournisseurs doivent marquer les contenus synthétiques de façon lisible par machine ; un code de pratique publié en juin 2026 et des lignes directrices de juillet 2026 en précisent la mise en œuvre (Mishcon de Reya, 2026). Aux États-Unis, plusieurs États encadrent déjà les répliques numériques, tandis qu'une loi fédérale reste à l'étude (tableau 3). Au Danemark, un projet de loi accordant à chacun un droit sur son visage et sa voix a été déposé le 8 octobre 2026 ; plusieurs publications l'ont présenté à tort comme adopté (The AI Musicpreneur, 2026).

Tableau 3

Principaux régimes visant les répliques numériques au 17 novembre 2026

Territoire	Texte	État	Exigence principale
Québec	C.c.Q. art. 35-36 ; Loi 25 ; LCCJTI art. 44-45	En vigueur	Consentement ; biométrie déclarée à la CAI
Union européenne	AI Act, art. 50	En vigueur depuis le 2 août 2026	Divulgation des hypertrucages ; marquage
États-Unis (fédéral)	NO FAKES Act (S.4591)	Approuvé en comité ; non adopté	Droit fédéral sur la voix et l'image
Californie	AB 2602 ; AB 1836	En vigueur depuis le 1er janv. 2025	Description précise des usages au contrat ; personnes décédées
New York	S7676B	En vigueur depuis le 1er janv. 2025	Même exigence contractuelle
Tennessee	ELVIS Act	En vigueur (2024)	Interdit l'imitation non consentie de la voix
Danemark	Projet de loi L 49	Déposé le 8 oct. 2026 ; non adopté	Droit sur son visage et sa voix (visé : 2027)

Note. Sources : Congrès des États-Unis (2026) ; Davis Wright Tremaine (2025) ; Mishcon de Reya (2026) ; Morrison Foerster (2025) ; Paubox (2026) ; The AI Musicpreneur (2026).

6.6 Normes professionnelles

Le milieu du cinéma a fixé une norme utile au monde de l'entreprise. La convention SAG-AFTRA de 2023 exige un consentement éclairé, clair et distinct — non enfoui dans les conditions générales —

pour créer et utiliser la réplique numérique d'un interprète ; elle prévoit une rémunération, et un consentement séparé pour chaque nouveau projet (Perkins Coie, 2024 ; SAG-AFTRA, 2023). Les lois de Californie et de New York exigent de même, pour être valides, que les clauses de réplique numérique décrivent de façon raisonnablement précise les usages prévus (Davis Wright Tremaine, 2025).

6.7 Provenance et prévention de la fraude

La norme C2PA (« Content Credentials ») permet d'attacher à une image, une vidéo ou un son un historique de provenance signé cryptographiquement ; elle compte plus de 6 000 membres et affiliés (TrueScreen, 2026), sa version 2.3 a étendu la provenance à la vidéo en direct et sa version 2 est en voie de normalisation ISO (PPC Land, 2026). La provenance certifie l'histoire d'un contenu, non sa véracité ; combinée au consentement et à des protocoles internes — par exemple, ne jamais autoriser un virement sur la seule foi d'une vidéo —, elle aide à distinguer le double autorisé de l'imitation frauduleuse illustrée par l'affaire Arup.

6.8 Enjeux éthiques

Dignité et contrôle. La personne doit pouvoir approuver chaque contenu, refuser un usage et retirer son consentement ; son double ne doit jamais lui faire dire ce qu'elle n'aurait pas dit.

Propriété et dépendance. Un pionnier des humains numériques, placé sous séquestre en février 2026, se réservait par contrat la propriété des avatars personnalisés de ses clients (BERI, 2026 ; UneeQ, 2026). La question « à qui appartient mon double et que devient-il si le fournisseur disparaît ? » doit être réglée au contrat.

Équité. Les biais observés dans la perception des visages de synthèse (Miller et al., 2023) imposent de vérifier que la qualité du double ne varie pas selon la couleur de peau, l'âge ou le genre de la personne.

Après la personne. Le décès, le départ de l'entreprise ou la fin du mandat doivent être prévus : la Californie encadre déjà les répliques de personnes décédées (AB 1836), et la convention SAG-AFTRA exige alors le consentement des ayants droit (Davis Wright Tremaine, 2025 ; SAG-AFTRA, 2023).

7. Marché et usages en entreprise

7.1 Estimations publiées

Les estimations de la taille du marché divergent d'un ordre de grandeur, principalement parce que les firmes ne définissent pas le marché de la même façon.[6] Nous les reproduisons à titre indicatif.

[6] Les montants sont en dollars américains, tels que publiés. Les périmètres diffèrent : certaines firmes incluent les influenceurs virtuels et le jeu vidéo, d'autres seulement les avatars d'entreprise.

Tableau 4

Estimations de marché publiées

Firme	Périmètre	Estimation publiée
Market Glass	Avatars humains numériques	14,1 G$ (2024) → 116,7 G$ (2030)
Technavio	Avatars humains numériques	7,47 G$ (2025) ; croissance annuelle de 43 % (2026-2030)
MarketsandMarkets	Avatars IA	0,80 G$ (2025) → 5,93 G$ (2032)

Note. Sources : Market Glass (2025) ; MarketsandMarkets (s.d.) ; Technavio (2026).

7.2 Signaux du marché

Deux signaux sont plus fiables que ces projections. D'une part, l'investissement : un chef de file britannique de la vidéo d'avatars a été valorisé à 4 milliards de dollars américains en janvier 2026 et oriente sa croissance vers des agents conversationnels (HLTH, 2026). D'autre part, la fragilité : un pionnier néo-zélandais des humains numériques a été placé sous séquestre en février 2026 (UneeQ, 2026). Le marché valorise la production à grande échelle, mais punit les modèles qui ne trouvent pas d'usages durables ; pour un client, la pérennité du fournisseur et la portabilité du double deviennent des critères d'achat.

7.3 Familles d'usage

Tableau 5

Usages d'entreprise du double numérique : valeur, risques et garde-fous

Usage	Valeur attendue	Risque principal	Garde-fou
Communication de la direction (résultats, annonces)	Présence du dirigeant sans tournage	Perte de crédibilité si le rendu est approximatif	Réalisme élevé ; mention explicite ; approbation
Borne d'accueil et de conseil (magasin, hôtel, salon)	Accueil continu, multilingue, avec présence sociale	Réponse erronée engageant l'entreprise	Sources approuvées ; abstention ; relais humain
Version web des réponses	Même service en ligne, sans calcul par visiteur	Contenus périmés	Banque validée et datée ; révision périodique
Réalité virtuelle (événements, lancements)	Effet de présence fort	Inconfort si le réalisme est moyen	Rendu soigné ; séances courtes
Vœux et messages personnalisés	Message nominatif à grande échelle	Impression de manipulation	Transparence ; texte approuvé par la personne
Formation	Geste montré par la bonne personne, dans chaque langue	Écart entre le geste et la procédure	Validation par l'expert métier

8. Méthode VP Capture

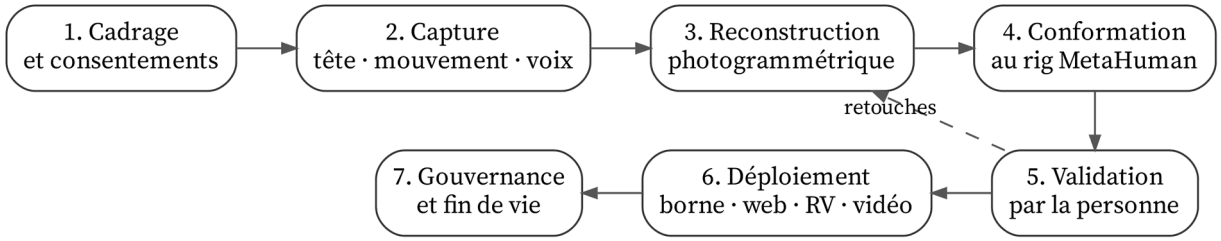
La méthode que nous appliquons traduit les constats des sections 4 à 6 en pratiques. Elle est décrite telle qu'elle est conçue ; ses mesures de performance feront l'objet d'une publication distincte (section 10). Elle s'organise en sept étapes (figure 1) et quatre exigences transversales : réalisme, fiabilité, consentement et gouvernance.

8.1 Cadrage et consentements

Avant toute capture, les usages prévus sont décrits avec précision et la personne signe deux consentements distincts : l'un pour son double numérique, l'autre pour sa voix. Chacun énumère les usages autorisés, peut être retiré en tout temps et prévoit l'effacement des données dans les 30 jours suivant un retrait. Cette pratique suit l'esprit de l'article 14 de la loi québécoise sur le secteur privé, de la convention SAG-AFTRA et des lois californienne et new-yorkaise (sections 6.2 et 6.6).

Figure 1

Chaîne de production du double numérique en sept étapes



Note. La validation par la personne peut renvoyer le modèle à l'étape de reconstruction (flèche en pointillé).

8.2 Capture

- **Tête et visage.** Colonnes de caméras à capteur Sony IMX477 montées sur pieds, déclenchées simultanément avec flash, pendant que la personne tourne sur un siège. Deux formules : Régulière (84 photographies) et Haute qualité (132 photographies, avec pause avant les expressions). L'approche multivue avec calibrage d'échelle suit les conditions qui, dans la littérature, approchent le millimètre (Giuliani et al., 2024).
- **Mouvement.** Capture sans combinaison ni marqueurs par plusieurs caméras calibrées, triangulée par un logiciel libre de cinématique multivue (Pagnon et al., 2022), pour la démarche, les gestes et la façon de saluer de la personne.
- **Voix.** Enregistrement guidé avec textes préparés et contrôle de chaque prise ; le clonage est traité sur nos propres serveurs plutôt que chez un tiers.

8.3 Reconstruction, conformation et validation

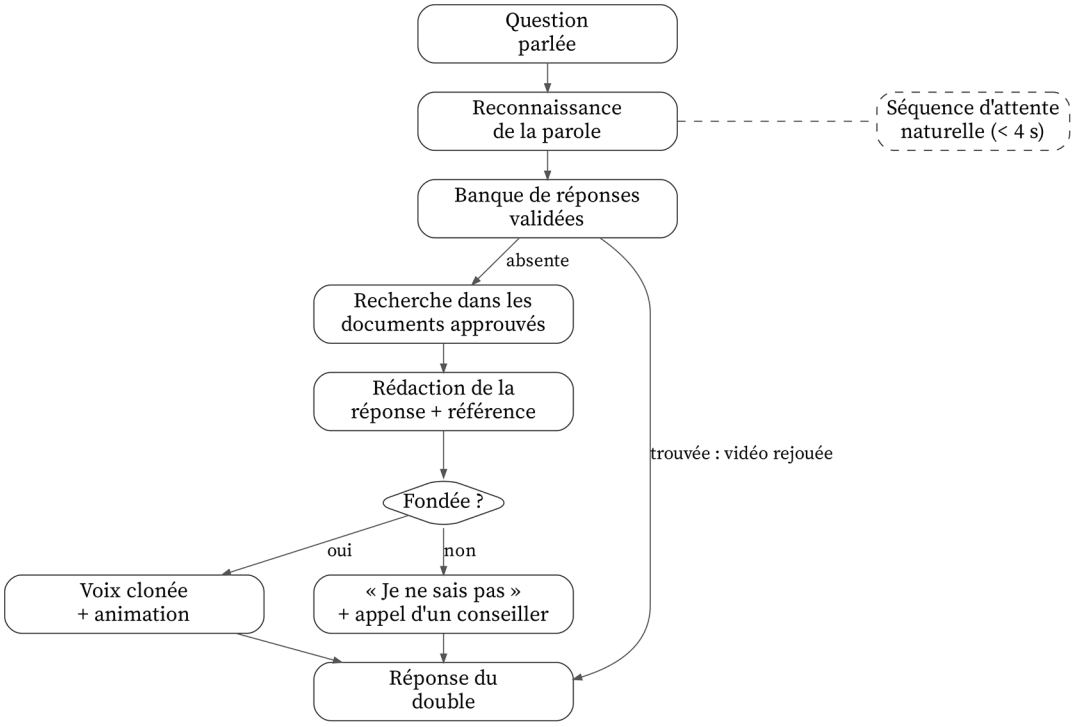
Le modèle reconstruit est conformé au rig MetaHuman, ce qui donne un personnage standardisé, animable et portable entre outils (section 4.2). L'objectif explicite est le réalisme élevé : un résultat

intermédiaire n'est pas livré, puisque c'est la zone que la littérature associe au malaise (Tao et al., 2025). La personne valide son double avant tout usage ; ses remarques peuvent renvoyer le modèle à l'étape de reconstruction. La qualité est vérifiée avec le même soin quels que soient la couleur de peau, l'âge ou le genre de la personne (section 6.8).

8.4 Fiabilité du double conversationnel

Figure 2

Architecture conversationnelle avec banque de réponses validées et abstention



Note. La séquence d'attente démarre dès la fin de la question, pour que l'attente perçue reste sous le seuil de 4 secondes rapporté par Maslych et al. (2025).

- **Abstention plutôt qu'invention.** Le personnage ne répond qu'à partir des documents approuvés du client, avec la référence officielle ; à défaut, il dit « je ne sais pas » et propose d'appeler un conseiller (Joren et al., 2025 ; Moffatt c. Air Canada, 2024).
- **Banque de réponses validées.** Les questions fréquentes sont vérifiées par le client, rendues une fois en qualité cinéma, puis rejouées. C'est aussi ce qui rend possible la version web sans calcul graphique par visiteur ; chaque nouvelle question est rendue une fois puis ajoutée à la banque.
- **Latence habitée.** Des séquences préparées d'avance — gestes de réflexion, déplacements — couvrent le temps de calcul, conformément aux remplissages naturels étudiés par Maslych et al. (2025).

8.5 Déploiement, transparence et fin de vie

- **Transparence.** Le public est informé qu'il s'adresse à un double numérique, conformément à l'esprit de l'article 50 de l'AI Act.
- **Approbation de chaque contenu.** La personne approuve chaque vidéo avant diffusion.
- **Sobriété des données à la borne.** Aucune reconnaissance faciale des visiteurs ; les photographies et vidéos du photomaton sont effacées automatiquement après 7 jours.
- **Fin de vie.** Départ, retrait du consentement ou fin du contrat déclenchent l'effacement dans le délai prévu.

9. Discussion

9.1 Réponses aux questions de recherche

QR1. La littérature récente ne confirme pas la crainte d'un réalisme excessif : le réalisme élevé favorise la confiance, et c'est l'à-peu-près qui la détruit (Tao et al., 2025 ; Baake et al., 2025). Le corollaire est éthique : puisque le public distingue mal le vrai du synthétique (Nightingale et Farid, 2022), l'obligation de transparence revient au producteur. **QR2.** La fiabilité repose moins sur l'intelligence du modèle que sur deux disciplines : savoir s'abstenir lorsque les sources manquent, et habiter l'attente par des comportements naturels. **QR3.** Malgré des régimes différents, le Québec, l'Union européenne et plusieurs États américains convergent vers un consentement écrit, précis, distinct et révocable, complété par la transparence envers le public.

9.2 Dix principes de bonne pratique

1. **Viser le réalisme élevé ou s'abstenir.** Le réalisme moyen est la zone de malaise.
2. **Capter avant de générer.** Une numérisation multivue calibrée de la vraie personne vaut mieux qu'une ressemblance approximative.
3. **Traiter la voix comme le visage.** Elle est aussi identifiante et exige le même consentement.
4. **Consentir par écrit, usage par usage.** Description précise des usages ; consentement séparé pour chaque nouveau projet.
5. **Garder la personne aux commandes.** Approbation de chaque contenu, retrait en tout temps, effacement dans un délai défini.
6. **Régler la propriété du double.** Le contrat dit à qui appartiennent les données et le personnage, et ce qu'il advient si le fournisseur disparaît.
7. **Préférer « je ne sais pas » à l'erreur.** Réponses fondées sur des sources approuvées, avec référence.
8. **Rester sous 4 secondes.** Et habiter l'attente par des comportements naturels, pas par une icône.

9. **Dire au public qu'il parle à un double.** La transparence protège la confiance et répond aux obligations émergentes.
10. **Rendre les contenus traçables.** La provenance signée aide à distinguer le double autorisé de l'hypertrucage.

9.3 Grille d'évaluation d'un fournisseur

Les principes précédents se traduisent en questions qu'une organisation peut poser à tout fournisseur de doubles numériques :

Tableau 6

Grille d'évaluation d'un fournisseur de doubles numériques

Domaine	Question à poser
Capture	Le double est-il construit à partir d'une capture multivue calibrée de la vraie personne ?
Réalisme	Puis-je voir le double en gros plan, en mouvement et en parole, avant de signer ?
Consentement	Les consentements image et voix sont-ils distincts, précis et révocables ?
Approbation	La personne approuve-t-elle chaque contenu avant diffusion ?
Propriété	À qui appartiennent les données et le personnage ? Que se passe-t-il si vous cessez vos activités ?
Données	Où les données du visage et de la voix sont-elles traitées et conservées, et combien de temps ?
Fiabilité	Comment le personnage s'abstient-il lorsqu'il ne sait pas ? Ses réponses citent-elles leur source ?
Latence	Quel est le délai de réponse mesuré, et comment l'attente est-elle comblée ?
Transparence	Comment le public est-il informé qu'il s'adresse à un double ?
Provenance	Les vidéos produites portent-elles une provenance vérifiable ?

10. Limites et programme de mesure

Cette revue comporte des limites. Elle est narrative et non systématique ; d'autres travaux pertinents ont pu échapper à notre recherche. Plusieurs études de perception portent sur des images fixes ou des vidéos courtes, en laboratoire ; leurs conclusions doivent être confirmées pour des personnages animés et conversationnels, en milieu réel et dans la durée. Les résultats sur la vidéo personnalisée proviennent du marketing grand public. La précision de capture citée provient d'une étude de faisabilité. Enfin, le droit évolue rapidement : plusieurs textes cités ne sont encore que des projets, et l'état décrit ici devra être mis à jour.

Pour combler ces lacunes, VP Capture Studio mesurera ses propres déploiements pilotes et en publiera les résultats selon les indicateurs du tableau 7.

Tableau 7

Indicateurs du programme de mesure

Indicateur	Définition	Méthode
Précision géométrique	Écart entre le double et une référence mesurée	Comparaison de nuages de points sur objet étalon et volontaires
Latence de bout en bout	Délai entre la fin de la question et le début de la réponse	Journalisation horodatée à la borne
Justesse	Part des réponses fondées, des abstentions et des erreurs	Revue d'un échantillon par le client
Confiance perçue	Expertise, intégrité et bienveillance attribuées	Questionnaire validé après l'interaction
Réception interne	Visionnement et réactions aux messages personnalisés	Statistiques d'envoi et sondage anonyme

11. Conclusion

La technologie du double numérique a mûri plus vite que les pratiques qui l'encadrent. Les outils sont puissants et de plus en plus ouverts ; le public distingue mal le vrai du synthétique ; et le droit, encore fragmenté, converge vers une exigence simple : le consentement éclairé de la personne et la transparence envers le public. Dans ce contexte, l'avantage ne revient pas à celui qui produit le plus vite, mais à celui qui produit le plus juste : un réalisme sans compromis, un personnage qui sait se taire lorsqu'il ne sait pas, et une personne réelle qui garde, du début à la fin, le contrôle de son image et de sa voix. C'est la méthode que nous appliquons, et celle que nous invitons les organisations à exiger de tout fournisseur.

Glossaire

Abstention. Refus motivé de répondre lorsque les sources disponibles ne permettent pas une réponse fondée.
Calibrage d'échelle. Opération qui donne au modèle 3D ses dimensions réelles à partir de distances connues.
Conformation. Ajustement d'un modèle numérisé à une topologie et un rig standards.
C2PA. Norme ouverte de provenance qui attache un historique signé à un contenu numérique.
FACS. Système de codage des actions faciales, qui décompose les expressions en unités d'action.
Génération augmentée par recherche. Architecture qui fournit au modèle de langage des passages documentaires au moment de répondre.
Photogrammétrie. Reconstruction 3D à partir de photographies prises sous plusieurs angles.
Rig. Squelette et commandes qui permettent d'animer un personnage.
Vallée dérangeante. Chute de l'affinité envers une figure presque, mais pas tout à fait, humaine.

Références

Animation World Network. (2025, septembre). *NVIDIA open sources Audio2Face animation model*. https://awn.com/news/nvidia-open-sources-audio2face-animation-model

Baake, J., Schmitt, J. B., & Metag, J. (2025). Balancing realism and trustworthiness: AI avatars in science communication. *Journal of Science Communication, 24*(02), A03. https://doi.org/10.22323/2.24020203

Barrington, S., Cooper, E. A., & Farid, H. (2025). People are poorly equipped to detect AI-powered voice clones. *Scientific Reports, 15*, 11004. https://doi.org/10.1038/s41598-025-94170-3

BERI. (2026). *AppDirect buys Soul Machines, which owns your custom avatar*. https://www.beri.net/article/appdirect-soul-machines-acquisition-receivership-custom-avatar-ownership-contract-terms

Business Today. (2025, 23 mai). *Tech CEOs use AI avatars for earnings calls*. https://www.businesstoday.in/technology/news/story/tech-ceos-use-ai-avatars-for-earnings-calls-signalling-a-shift-in-corporate-communication-477455-2025-05-23

CASRAI. (2026). *AIDA (Bill C-27): What it would have required*. https://casrai.org/guides/aida-bill-c27-what-it-would-have-required

Cassell, J., Sullivan, J., Prevost, S., & Churchill, E. (Éds). (2000). *Embodied conversational agents*. MIT Press.

CFO Dive. (2024). *Scammers siphon $25M from engineering firm Arup via AI deepfake CFO*. https://www.cfodive.com/news/scammers-siphon-25m-from-engineering-firm-arup-via-ai-deepfake-cfo/716501/

Charte des droits et libertés de la personne, RLRQ c C-12, art. 5.

Code civil du Québec, RLRQ c CCQ-1991, art. 35-36.

Congrès des États-Unis. (2026). *S.4591 — NO FAKES Act of 2026*. https://www.govinfo.gov/content/pkg/BILLS-119s4591is/xhtml/BILLS-119s4591is.html

Davis Wright Tremaine. (2025). *Lights, camera, legislation: Are your entertainment contracts AI ready?* https://dwt.com/insights/2025/03/state-laws-regulating-ai-in-entertainment-industry

Davis+Gilbert. (2024). *California passes new legislation prohibiting unauthorized AI replicas*. https://www.dglaw.com/california-passes-new-legislation-prohibiting-unauthorized-ai-replicas/

Dennis, M., & Kizer, T. H. (2025). *Lights, camera, engagement! The role of personalized video as a marketing strategy in capturing consumer attention*. https://scholarship.rollins.edu/as_facpub/420

Dezeen. (2024, 17 mai). *Arup victim of multimillion-dollar deepfake video scam in Hong Kong*. https://www.dezeen.com/2024/05/17/arup-victim-deepfake-video-scam/

Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System*. Consulting Psychologists Press.

Epic Games. (s.d.). *Unreal Engine — Foire aux questions sur les licences*. https://gateway.unrealengine.com/en-US/faq

Epic Games. (2026a). *MetaHuman news*. https://www.metahuman.com/en-US/news

Epic Games. (2026b). *MetaHuman 5.8 release notes in Unreal Engine*. https://dev.epicgames.com/documentation/metahuman/metahuman-5-8-release-notes-in-unreal-engine

Epic Games. (2026c). *OpenRigLogic* [Logiciel, licence MIT]. GitHub. https://github.com/EpicGames/openriglogic

Epic Games. (2026d). *Unreal Engine 5.8 is now available*. https://www.unrealengine.com/en-US/news/unreal-engine-5-8-is-now-available

Giuliani, S., Tosti, F., Lopes, P., Ciampini, C., & Nardinocchi, C. (2024). Design of a multi-view photogrammetric system based on low cost cameras for the 3D forensic face recognition. *ISPRS Archives, XLVIII-2/W8-2024*, 177–183. https://doi.org/10.5194/isprs-archives-XLVIII-2-W8-2024-177-2024

Gowling WLG. (s.d.). *Biométrie : cadre juridique québécois*. https://gowlingwlg.com/getmedia/7ec810e8-6b8e-4d3d-bda6-ee3b568d60dd/Gowling-WLG-Quebec-Biometrie-Legal-Framework-Fr-v2.pdf.xml

HLTH. (2026, 27 janvier). *Synthesia raises $200M, with $4B valuation*. https://hlth.com/insights/news/synthesia-raises-200m-with-4b-valuation-2026-01-27

InformalScience. (2025). *The impact of a conversational agent avatar on the experience of museum visitors* [Évaluation]. https://informalscience.org/evaluation/the-impact-of-a-conversational-agent-avatar-on-the-experience-of-museum-visitors/

Joren, H., Zhang, J., Ferng, C.-S., et al. (2025). Sufficient context: A new lens on retrieval augmented generation systems. *ICLR 2025*. https://arxiv.org/abs/2411.06037

Kapoor, A., & Kumar, M. (2025). *Frontiers: Generative AI and personalized video advertisements*. https://ide.mit.edu/insights/personalized-ai-video-ads

Kerbl, B., Kopanas, G., Leimkühler, T., & Drettakis, G. (2023). 3D Gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics, 42*(4), 139. https://doi.org/10.1145/3592433

Kersten, T. P., Timm, F., & Zobel, K. (2024). Development of a photogrammetric 3D measurement system for small objects using Raspberry Pi cameras as low-cost sensors. *ISPRS Archives, XLVIII-2/W8-2024*, 219–225. https://doi.org/10.5194/isprs-archives-XLVIII-2-W8-2024-219-2024

Lewis, P., Perez, E., Piktus, A., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems, 33*. https://arxiv.org/abs/2005.11401

Loi concernant le cadre juridique des technologies de l'information, RLRQ c C-1.1, art. 44-45.

Loi sur la protection des renseignements personnels dans le secteur privé, RLRQ c P-39.1, art. 14.

Market Glass. (2025, novembre). *Digital human avatars: Global market report*. MarketResearch.com. https://www.marketresearch.com/Market-Glass-Inc-v1039/Digital-Human-Avatars-42950110

MarketsandMarkets. (s.d.). *AI avatar market*. https://www.marketsandmarkets.com/Market-Reports/digital-human-market-212510229.html

Maslych, M., Katebi, M., Lee, C., et al. (2025). Mitigating response delays in free-form conversations with LLM-powered intelligent virtual agents. *CUI '25*. ACM. https://arxiv.org/abs/2507.22352

Miller, E. J., Steward, B. A., Witkower, Z., Sutherland, C. A. M., Krumhuber, E. G., & Dawel, A. (2023). AI hyperrealism: Why AI faces are perceived as more real than human ones. *Psychological Science, 34*(12), 1390–1403. https://doi.org/10.1177/09567976231207095

Mishcon de Reya. (2026). *AI Act transparency obligations: Code of practice and draft guidelines*. https://www.mishcon.com/news/ai-act-transparency-obligations-code-of-practice-and-draft-guidelines

Moffatt v. Air Canada, 2024 BCCRT 149 (CanLII). https://www.canlii.org/en/bc/bccrt/doc/2024/2024bccrt149/2024bccrt149.html

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [From the field]. *IEEE Robotics & Automation Magazine, 19*(2), 98–100. https://doi.org/10.1109/MRA.2012.2192811

Morrison Foerster. (2025). *Digital avatars deep dive series: Navigating the legal and regulatory landscape in 2025*. https://www.mofo.com/pdf/resources/insights/250922-digital-avatars-deep-dive-series-navigating

Nightingale, S. J., & Farid, H. (2022). AI-synthesized faces are indistinguishable from real faces and more trustworthy. *Proceedings of the National Academy of Sciences, 119*(8), e2120481119. https://doi.org/10.1073/pnas.2120481119

Pagnon, D., Domalain, M., & Reveret, L. (2022). Pose2Sim: An open-source Python package for multiview markerless kinematics. *Journal of Open Source Software, 7*(77), 4362. https://doi.org/10.21105/joss.04362

Paubox. (2026). *Senate committee advances NO FAKES Act*. https://www.paubox.com/blog/senate-committee-advances-no-fakes-act

Perkins Coie. (2024). *Generative AI in movies and TV: How the 2023 SAG-AFTRA and WGA contracts address generative AI*. https://perkinscoie.com/insights/blog/generative-ai-movies-and-tv-how-2023-sag-aftra-and-wga-contracts-address-generative

PPC Land. (2026). *C2PA*. https://ppc.land/c2pa/

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press.

SAG-AFTRA. (2023). *Digital replicas 101: What you need to know about the 2023 TV/Theatrical contracts*. https://sagaftra.org/files/sa_documents/DigitalReplicas.pdf

Saito, S., Schwartz, G., Simon, T., Li, J., & Nam, G. (2024). Relightable Gaussian codec avatars. *CVPR 2024*. https://arxiv.org/abs/2312.03704

Schönberger, J. L., & Frahm, J.-M. (2016). Structure-from-motion revisited. *CVPR 2016*, 4104–4113. https://doi.org/10.1109/CVPR.2016.445

Tao, Z., Liu, Y., Qiu, J., & Li, S. (2025). Impact of virtual avatar appearance realism on perceptual interaction experience: A network meta-analysis. *Frontiers in Psychology, 16*, 1624975. https://doi.org/10.3389/fpsyg.2025.1624975

TechCrunch. (2025). *After Klarna, Zoom's CEO also uses an AI avatar on quarterly call*. https://techcrunch.com/?p=3011302

Technavio. (2026, mai). *Digital human avatar market analysis, size, and forecast 2026-2030*. https://www.technavio.com/report/digital-human-avatar-market-industry-analysis

TechSpot. (2025). *The CEOs of Zoom and Klarna have presented earnings calls using AI avatars*. https://www.techspot.com/news/108068-ceos-zoom-klarna-have-presented-earnings-calls-using.html

Thacker, J. (2025, 4 juin). *You can now sell MetaHumans, or use them in Unity or Godot*. CG Channel. https://www.cgchannel.com/2025/06/you-can-now-sell-metahumans-or-use-them-in-unity-or-godot/

The AI Musicpreneur. (2026). *Denmark digital replica bill L 49*. https://www.aimusicpreneur.com/ai-music-news/denmark-digital-replica-bill-ai-voice-likeness-consent-2027/

Torys. (2025, avril). *La CAI et les renseignements biométriques*. https://www.torys.com/fr-ca/our-latest-thinking/publications/2025/04/cai-renseignements-biometriques

Traum, D., Aggarwal, P., Artstein, R., Foutz, S., Gerten, J., Katsamanis, A., Leuski, A., Noren, D., & Swartout, W. (2012). Ada and Grace: Direct interaction with museum visitors. Dans *Intelligent Virtual Agents (IVA 2012)*. Springer. https://doi.org/10.1007/978-3-642-33197-8_25

TrueScreen. (2026). *C2PA standard: History, promises and structural limitations*. https://truescreen.io/articles/c2pa-standard-history-limitations/

UneeQ. (2026, 9 février). *UneeQ vs. Soul Machines: The best digital human platforms compared*. https://www.digitalhumans.com/blog/uneeq-vs-soul-machines-best-digital-human-platforms-compared

Usercentrics. (2026). *Consentement et Loi 25*. https://usercentrics.com/knowledge-hub/consentement-loi-25/

Zhang, D., Liu, Y., Lin, L., Zhu, Y., Chen, K., Qin, M., Li, Y., & Wang, H. (2025). HRAvatar: High-quality and relightable Gaussian head avatar. *CVPR 2025*, 26285–26296. https://doi.org/10.1109/CVPR52734.2025.02448

Zolynski, C. (2026). Deepfakes and transparency obligations. *Columbia Journal of Law & the Arts, 49*(4), 773.