

RAPPORT FINAL

Pilote Lotus & Cali

Ce que huit mois de production IA
révèlent sur l'avenir de l'animation
jeunesse

Sommaire exécutif

L'ambition était simple à formuler, redoutable à exécuter : confier à l'IA générative la production complète d'un épisode de *Lotus & Cali*.

D'octobre 2024 à juin 2025, Vooban et Télé-Québec ont mené ce pilote. Pas pour livrer un produit fini prêt à diffuser. Pour répondre à une question que personne au Québec n'avait encore testée dans des conditions réelles : **jusqu'où l'IA peut-elle aller dans une chaîne de production audiovisuelle, et où s'arrête-t-elle?**

Le projet a été soutenu par le ministère de la Culture du Québec. Budget réel : environ **78 K\$** sur une enveloppe de 120 K\$. L'indicateur de performance retenu : le nombre d'apprentissages générés. Il a été largement atteint.

Mais au-delà des chiffres, c'est la nature de ces apprentissages qui donne au pilote sa valeur durable. Ce que l'IA sait déjà faire. Ce qu'elle ne sait pas encore faire. Et comment penser différemment la prochaine itération. Le présent rapport rend compte de ces apprentissages du point de vue de Vooban.

Contexte et mandat

Télé-Québec a choisi *Lotus & Cali*, série éducative destinée aux 4-7 ans en développement depuis 2017, pour une raison précise : elle offrait un comparable exigeant. Une production existante de haute qualité, contre laquelle mesurer ce que l'IA pouvait accomplir. Et ce qu'elle ne pouvait pas.

Vooban, firme québécoise spécialisée en IA, a piloté la réalisation. La collaboration était étroite : équipes de contenu, de production et service juridique de Télé-Québec, tous impliqués. Un chercheur de l'UQAM a assuré la documentation académique en parallèle, offrant un regard externe complémentaire aux observations de terrain de Vooban et à la documentation interne de Télé-Québec.

Méthodologie

Deux logiques de production face à face

En production traditionnelle, un épisode d'animation traverse trois grandes phases : pré-production, production, post-production. Plus de quinze rôles spécialisés. L'échéancier est linéaire : chaque livrable conditionne le suivant. On ne s'engage pas dans les dessins de qualité finale tant que le scénario et le scénarimage ne sont pas verrouillés. Chaque itération coûte cher en temps humain.

Le flux de travail IA développé par Vooban compresse cette chaîne en huit étapes, orchestrées autour d'une dizaine d'outils spécialisés.

La pré-production s'appuie sur des modèles de langage (Claude, ChatGPT) pour l'idéation et l'écriture. Scenario.com sert à entraîner un modèle au style visuel de l'émission, puis à générer personnages et décors. Le montage se fait dans DaVinci Resolve. Flux Kontext et Nano Banana assurent la cohérence visuelle des arrière-plans et des personnages.

Méthodologie

L'animation procède par blocs de cinq secondes, avec un choix d'outils selon la qualité visée : Veo 3 pour le haut de gamme, Seedance pour la production courante. ElevenLabs génère les voix synthétiques. HeyGen prend en charge la synchronisation labiale, après avoir rapidement écarté Sync.so, nettement inférieur. La post-production boucle avec colorisation et ajustement de la cadence à 15 images par seconde, pour retrouver le rendu d'animation classique.

Un exemple révélateur : l'équipe avait d'abord tenté d'animer personnages et décors ensemble dans un même plan. On a rapidement adopté la technique du fond vert, ironiquement empruntée au cinéma traditionnel. Créer les personnages séparément, les détourner, puis les intégrer sur les décors dans DaVinci Resolve. L'IA avait beau être nouvelle, les contraintes de composition d'images restaient les mêmes.

La méthodologie relevait du développement logiciel, pas de l'animation. Tester, apprendre, tester, apprendre. On pouvait recommencer sans pénalité de coût significative. Ce changement de posture a été un apprentissage en soi.

Workflow de réalisation d'épisode animé assisté par IA

Phase	Étapes	Outils
1. Création du personnage	Idéation et génération de plusieurs concepts, sélection du personnage (nom, description), création du backstory, entraînement d'un modèle au style L&C, génération visuelle itérative, génération finale	LLMs, Scenario.com
2. Concept et scénario	Développement du concept d'épisode (scénario, décors, personnages, dialogues, effets sonores, musique), développement des personnages, écriture des scénarios	Claude, ChatGPT
3. Montage initial	Création du projet de montage, création des arrière-plans, ajustement des résolutions et versions multiples, raffinement des décors	DaVinci Resolve, Scenario.com, Flux Kontext, Nano Banana
4. Production visuelle	Placement des dialogues dans le montage, création des postures des personnages, ajustement de résolution, nettoyage des postures (fond vert)	Scenario.com, Flux Kontext, Nano Banana, Photoshop
5. Animation	Animation par blocs de 5 secondes, sélection du premier et dernier frame, génération de l'animation, insertion des images d'animation labiale	Veo 3 (haute qualité), Seedance (courante), Kling 3.0
6. Montage des blocs	Assemblage des blocs animés, suppression des fonds verts, insertion des arrière-plans	DaVinci Resolve
7. Audio et voix	Création et sélection des voix synthétiques, création du dialogue audio, réinsertion dans le montage, synchronisation labiale manuelle	Fish Audio S2 Pro, DaVinci Resolve
8. Édition finale	Synchronisation labiale automatisée, export image et son séparés, réassemblage, ajustement du FPS (30→15), colorisation, ajout de texture	Wan 2.1 FLF2V, DaVinci Resolve

Ce que le pilote a révélé

L'écriture : quand le probable est l'ennemi du créatif

Le premier scénario complet généré par IA a obtenu 20 % de l'équipe de contenu de Télé-Québec. Il aurait été refusé par n'importe quel producteur. Les émotions se succédaient sans logique. Le ton était mécanique. Les personnages manquaient de profondeur. Pour des enfants de 4 à 7 ans, rien ne fonctionnait.

L'équipe a pivoté vers l'écriture scène par scène, avec des règles éditoriales injectées progressivement dans les instructions du modèle. La vitesse de génération restait impressionnante : cinquante propositions en une minute. Mais d'atelier en atelier, un schéma récurrent s'est imposé. Chaque nouvelle règle corrigeait un problème et en créait un autre. On avançait d'un pas, on reculait d'un demi-pas.

Constat clé

Un modèle de langage prédit le mot le plus probable. Par définition, c'est l'inverse de la créativité. Or l'humour et l'émotion en contenu jeunesse exigent précisément de sortir du probable.

L'IA accélère l'idéation. Elle peut servir de partenaire de remue-méninges redoutable. Mais le jugement éditorial (savoir quand un personnage sonne vrai, quand une émotion atterrit) reste un territoire humain.

Le visuel : volume contre précision

Au départ, Scenario.com a permis d'entraîner un modèle dédié au style visuel de l'émission. Personnages et décors générés avec une cohérence de base. Par la suite, l'arrivée de Flux Kontext et Nano Banana a changé la donne : on pouvait modifier directement les dessins originaux de Lotus et Cali. Nouveaux décors, vêtements, postures, mouvements. Sans phase d'entraînement.

L'IA excelle en volume brut. Des centaines de propositions en quelques clics, autant pour les décors que pour les variantes de personnages. Mais le tri restait entièrement humain. Sur ces centaines de résultats, l'équipe en retenait une poignée qui s'approchaient du rendu habituel de la série.

La création de l'antagoniste Chut a nécessité **quatre itérations majeures** sans jamais atteindre le standard d'un animateur professionnel. Les personnages générés restaient trop semblables entre eux. Il leur manquait cette singularité qu'on reconnaît au premier coup d'œil dans une bonne conception de personnage. Le style particulier de *Lotus & Cali* (du 2D sans contours, avec du grain et des textures subtiles) s'est avéré particulièrement difficile à reproduire. Et plus le résultat s'approchait de l'original, plus les écarts devenaient flagrants.

L'animation souffrait d'un tempo uniformément lent. Pas de saccades, pas de lignes de mouvement. Tout ce qui donne vie aux vrais épisodes manquait. Des accessoires disparaissaient entre les plans. Des textures se perdaient d'une scène à l'autre.

La voix : le dernier bastion de l'artisanat

Réplica produisait un résultat très robotique en français. Nettement meilleur en anglais. Ça confirmait un retard marqué des modèles vocaux francophones et l'absence quasi totale de voix avec accent québécois. Pour un diffuseur public dont le mandat est ancré dans la culture d'ici, c'est une lacune bloquante.

ElevenLabs offrait une qualité supérieure, mais chaque réplique devait être générée dix à vingt fois pour en retenir une acceptable. Pour un épisode de soixante répliques, ça devient vite lourd.

Constat clé

Les voix jeunesse sont un art à part entière : incarner un enfant de façon crédible, moduler l'énergie, le rythme, l'émotion d'une ligne à l'autre. La synthèse vocale actuelle ne sait pas encore le faire. De tous les aspects testés, la voix est celui où l'écart entre l'IA et l'humain reste le plus grand.

Le facteur humain : l'apprentissage le plus transversal

Au fil du pilote, un constat s'est imposé avec de plus en plus de force : **ça prend beaucoup d'humains pour faire fonctionner l'IA.**

La production animée est une aventure collective. La friction créative entre personnes (les désaccords productifs, les rebonds d'idées, les surprises qui naissent du contact entre points de vue différents) génère la qualité finale. L'IA peut offrir des résultats heureux par hasard. Ce n'est pas la même chose qu'un élan créatif intentionnel qui naît de la collaboration.

Ce constat ne diminue pas la valeur de l'IA. Il la recadre. L'outil est puissant quand il amplifie l'humain. Il est faible quand il tente de le remplacer.

Gouvernance et éthique

Le test de transparence

En cours de projet, l'équipe Vooban a identifié un enjeu que personne n'avait anticipé dans son ampleur : plusieurs outils d'IA générative réutilisaient les données soumises pour réentraîner leurs propres modèles. Y compris les images des personnages de Télé-Québec.

Une analyse de sécurité exhaustive a été menée sur chaque service utilisé. Rétention de données, localisation des serveurs, politiques de réutilisation, risques liés aux sous-traitants tiers. Tout a été passé au crible.

Plutôt que de minimiser le problème, Vooban a choisi la **transparence totale**. Les résultats, validés par le responsable de la sécurité de l'information de Vooban, ont été présentés tels quels à Télé-Québec. Après avoir soupesé les risques face aux bénéfices d'apprentissage, Télé-Québec a choisi de poursuivre. Une décision éclairée plutôt qu'un risque ignoré.

La même approche a guidé les relations avec le milieu artistique. L'Union des artistes, informée dès le départ de l'utilisation de voix entièrement synthétiques, a donné son accord à la mi-avril 2025. Condition : qu'aucune voix existante ne soit utilisée comme base. La SARTEC et la SPACQ ont aussi été informées du projet.

Leçon de gouvernance

Dans un domaine où la méfiance envers l'IA est légitime et répandue, le choix d'impliquer les parties prenantes dès le début, plutôt que de demander pardon après, s'est avéré être le bon.

L'accueil interne : la curiosité plutôt que la crainte

Un apprentissage inattendu du pilote : la réaction des équipes internes de Télé-Québec. Plutôt que la résistance qu'on observe souvent face à l'IA dans les milieux créatifs, l'accueil a été marqué par la curiosité. Les employés voyaient dans l'IA de nouvelles possibilités, pas une menace à leur métier.

Le fait d'avoir impliqué les équipes dès le départ, de leur avoir donné un rôle actif dans l'évaluation et les orientations, a dissipé les appréhensions. Les artisans qui ont participé au pilote ont d'ailleurs exprimé le souhait de recommencer. Avec les apprentissages accumulés, ils savaient où investir leur énergie. Et où ne pas en perdre.

Rayonnement

Un projet qui a dépassé ses murs

L'indicateur de performance retenu pour ce pilote (le nombre d'apprentissages) ne se mesure pas uniquement dans les pages qui précèdent. Il se mesure aussi dans les portes qu'il a ouvertes.

Le projet a généré près d'une **quinzaine de mentions** dans les médias québécois, dont Radio-Canada et Les Connecteurs. Présentations dans des conférences majeures : ALL IN, Interface, Impact IA, Entertain-AI. Présentation à l'international, devant des radiodiffuseurs publics à Berne. Des podcasts et des entrevues ont prolongé la conversation.

Pour Télé-Québec, le pilote a fait plus que tester une technologie. Il a positionné l'organisation comme un acteur proactif et crédible de l'IA en médias. Invitée à des tables où elle n'aurait pas été considérée auparavant.

Recommandations

1. Ancrer le modèle hybride humain-IA comme norme de production

Le pilote a confirmé une chose clairement : l'IA ne porte pas une chaîne de production toute seule. Elle amplifie l'humain. Elle ne le remplace pas.

Une prochaine itération devrait partir d'un scénario écrit par des auteurs en collaboration avec l'IA, de voix interprétées par des comédiens, et d'une couche visuelle copilotée entre animateurs et outils génératifs.

Phénomène clé

Quand le texte est écrit par un humain et la voix interprétée par un comédien, le spectateur pardonne les imperfections visuelles générées par l'IA. L'intention humaine dans le récit et la voix crée un ancrage émotionnel qui compense les écarts d'image.

La question n'est plus « est-ce que l'IA va entrer dans la production? ». C'est « qui va le faire bien, et qui va se faire dépasser? »

2. Concevoir à partir de l'IA, pas autour d'elle

Le réflexe naturel : insérer l'IA dans la chaîne existante, comme un outil de plus dans la boîte. Le pilote démontre le contraire. La vraie valeur émerge quand on repense la création en fonction des capacités de l'IA. Explorer des formats, des univers visuels et des volumes de contenu qui seraient inaccessibles en production traditionnelle.

2. Concevoir à partir de l'IA, pas autour d'elle

D'autres l'ont compris. En France, le studio Animaj, natif de l'IA et spécialisé en contenu jeunesse, a levé **75 millions d'euros**. Il produit un épisode en cinq semaines contre vingt en production classique. 242 millions de téléspectateurs par mois. Google figure parmi ses investisseurs.

Au Royaume-Uni, la BBC a créé un laboratoire dédié à la création jeunesse par IA, dirigé par une ancienne de Disney. Netflix a confirmé sa première scène générée par IA, complétée dix fois plus rapidement que par la méthode traditionnelle.

Aucun de ces acteurs n'a greffé l'IA sur un processus existant. Chacun a rebâti son processus autour de l'IA. La différence est fondamentale.

3. Capitaliser sur les acquis du pilote, qui sont devenus la norme

La génération de décors, l'accélération des scénarimages et l'idéation assistée ont prouvé leur valeur pendant le pilote. Mais un fait nouveau donne à ces résultats une portée inattendue.

Fait marquant

Le flux de travail que l'équipe Vooban a construit par essais et erreurs est aujourd'hui reconnu comme la méthode de référence par la communauté internationale de production animée assistée par IA.

Les studios qui ont adopté des flux de travail IA matures rapportent des réductions de coûts de **20 à 30 %**. Deux tiers des scénaristes professionnels utilisent désormais des outils d'IA dans leur pratique.

3. Capitaliser sur les acquis du pilote

Télé-Québec a déjà investi publiquement dans ce savoir-faire. Résultats présentés en conférences nationales et internationales. Syndicats impliqués. Couverture de Radio-Canada et d'une quinzaine de médias. Les acquis existent. Les fondations sont posées.

Le pilote a aussi permis de comprendre pourquoi certains résultats étaient en deçà des attentes. Le tempo uniformément lent de l'animation, par exemple, n'était pas un défaut des outils. C'est un compromis fondamental : plus on pousse la cohérence visuelle d'un personnage (en augmentant le poids de référence), plus le mouvement se fige. C'est un arbitrage, pas une limitation.

Autre apprentissage concret : la qualité des images de référence fournies aux modèles influençait directement la cohérence des résultats. Souvent davantage que le choix du modèle lui-même. Ces diagnostics précis transforment les difficultés du pilote en paramètres maîtrisables pour la suite.

4. Déployer des équipes d'agents spécialisés pour la scénarisation

Les lacunes en scénarisation observées durant le pilote (incohérence émotionnelle, personnages unidimensionnels, prévisibilité narrative) découlaient d'une cause identifiable. Un seul modèle généraliste tentait de jouer tous les rôles à la fois. Chaque nouvelle règle injectée dans ses instructions en déréglait une autre.

La solution existe désormais en production, pas en théorie.

4. Déployer des équipes d'agents spécialisés

Claude Code d'Anthropic permet aujourd'hui de créer des extensions spécialisées (appelées compétences) et de les orchestrer en équipes coordonnées. Un module veille à la cohérence émotionnelle. Un autre maintient la fidélité aux personnages : tempérament, vocabulaire, relations. Un module mémoire conserve les règles de l'univers de *Lotus & Cali* et l'historique des épisodes. Un module pédagogie valide les objectifs éducatifs pour les 4 à 7 ans. Un module critique joue le rôle du chef de contenu. Ces spécialistes itèrent entre eux, écrivent, critiquent, révisent, reproduisant la dynamique d'une salle d'écriture professionnelle.

Le problème central du pilote (ajouter une contrainte qui en brisait une autre) est résolu par conception. Ce type d'orchestration était possible auparavant, mais exigeait un système sur mesure, coûteux à développer et à maintenir. Aujourd'hui, il se configure en quelques heures.

5. Ajuster la sélection d'outils à la lumière de l'évolution récente

Depuis la fin du pilote, l'écosystème a évolué de façon significative. Certaines recommandations initiales doivent être corrigées.

Synchronisation labiale et animation contrôlée. Wan 2.1 FLF2V offre une solution solide. Le contrôle par première et dernière image permet de diriger précisément le mouvement. Le modèle est ouvert. Aucun risque lié à la propriété intellectuelle.

Voix. ElevenLabs interdit explicitement la génération de voix d'enfants dans ses conditions d'utilisation. Pour une série jeunesse destinée aux 4 à 7 ans, c'est bloquant.

5. Ajuster la sélection d'outils

Fish Audio S2 Pro est la solution de remplacement : 48 balises de contrôle émotionnel, clonage vocal à partir de dix à quinze secondes d'enregistrement, aucune restriction sur les voix d'enfants, coût inférieur de 70 %.

Il faut noter qu'aucune solution de synthèse vocale avec accent québécois natif n'existe sur le marché. Le clonage de la voix d'un comédien québécois demeure l'approche la plus viable pour préserver l'identité sonore de Télé-Québec. Cette avenue reste encadrée par les ententes avec l'Union des artistes, qui en balisent l'utilisation. C'est un espace à explorer conjointement.

Génération vidéo. Seedance 2.0 présente des contraintes non documentées dans sa mise en marché. Un filtre de sécurité trop agressif qui bloque par erreur des contenus inoffensifs, notamment les visages humains réalistes (contrainte liée à la réglementation chinoise de ByteDance). Coût réel : environ 2,50 \$ par séquence de quinze secondes. Temps de génération lents.

Kling 3.0, en revanche, introduit l'ancrage de l'identité visuelle d'un personnage sur six plans consécutifs. Ça répond directement au problème de cohérence inter-plans qui avait contraint le pilote.

Leçon stratégique

La fermeture de Sora (OpenAI) le 24 mars 2026, après des coûts d'exploitation d'un million de dollars par jour, illustre un risque que le pilote avait déjà identifié : la dépendance à un seul outil est une vulnérabilité. La stratégie de diversification adoptée par l'équipe Vooban, avec une dizaine d'outils interchangeables, s'avère non seulement prudente mais presciente.

Conclusion

Ce pilote a produit quelque chose qui ne figurait dans aucun livrable prévu : un savoir-faire opérationnel en production animée assistée par IA.

Testé dans les conditions réelles d'un diffuseur public. Avec les contraintes d'un contenu jeunesse, le regard d'un chercheur universitaire et l'accord des partenaires syndicaux.

Ce savoir-faire n'existe pas ailleurs au Québec. Il existe dans très peu d'endroits au monde.

Les équipes de Télé-Québec qui ont participé au pilote veulent recommencer. Les artisans qui ont collaboré savent où investir leur énergie et où ne pas en perdre. Les erreurs ont été documentées avec autant de soin que les réussites. Le flux de travail que le pilote a construit est devenu la référence internationale. Le positionnement public de Télé-Québec comme acteur crédible de l'IA en production a ouvert des portes qui n'existaient pas.

Pendant ce temps, un studio français conçu autour de l'IA produit déjà à l'échelle industrielle. La BBC a ouvert un laboratoire dédié. Netflix accélère. Le CRTC a posé le cadre réglementaire. Les syndicats ont signé les ententes. Les outils qui manquaient il y a dix mois existent aujourd'hui. Ceux d'aujourd'hui seront dépassés dans six mois.

Tout ce que le pilote a bâti (les apprentissages, les relations, le positionnement, les compétences) constitue un actif. Mais c'est un actif particulier : dans un domaine qui se transforme à ce rythme, ce qui n'est pas exploité ne reste pas en place. Il se dévalue.

Les fondations sont là. Les équipes sont prêtes. Le moment est maintenant.

De la part de toute
l'équipe, merci!

vooban.com

allo@vooban.com

Québec - Montréal - Toronto

1 844 800-0027