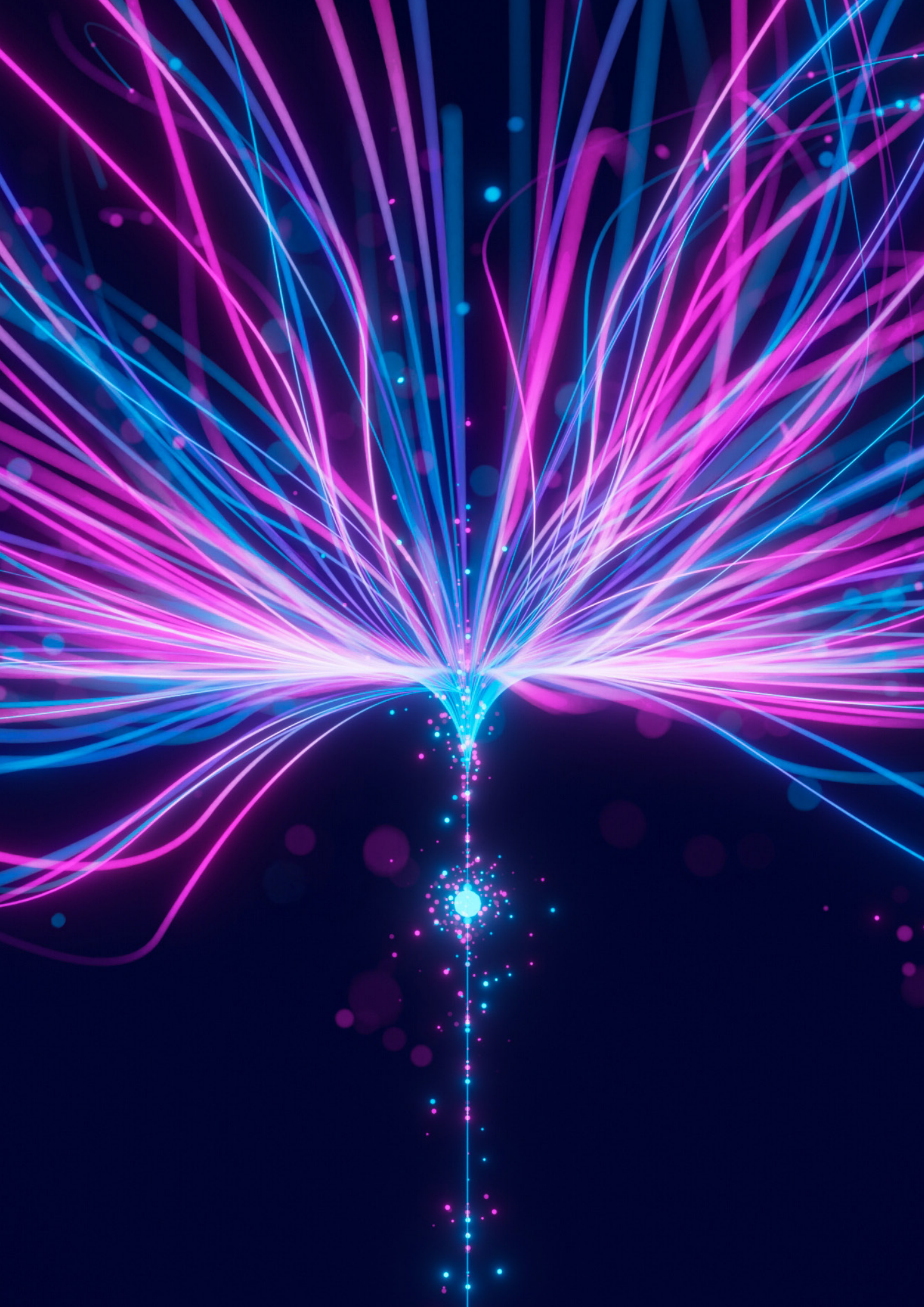
LIVRE BLANC

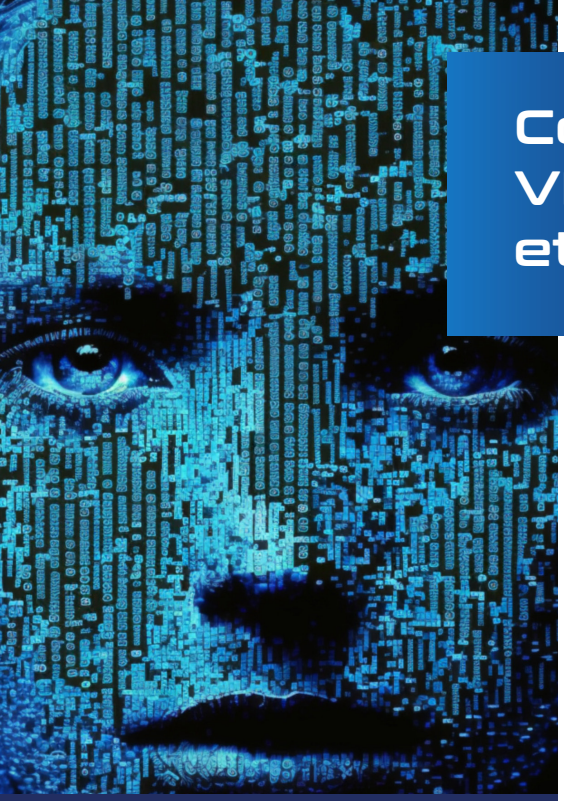


STOP AU VRAC NUMÉRIQUE !

Spark ARCHIVES

Une société de KLEE GROUP





Comment la solution ZERO VRAC libère l'espace, le temps et réduit l'empreinte carbone

« Le vrac numérique n'est pas une fatalité mais une décision différée. Décider, c'est créer de la valeur. »

Chiffres clés 2025

- **80 %** des données d'entreprise sont non structurées.
- **40 %** de doublons relevés sur un périmètre avant-vente en mai 2025.
- **1 To** inutile coûte 920 € et 25 kg CO_2 par an (ADEME).
- **12 %** seulement des fichiers conservés seront finalement réutilisés.
- **1 projet sur 5** subit un retard en raison d'une recherche documentaire infructueuse.

Vos baies de disques, vos partages réseau et vos abonnements cloud gonflent en silence. Depuis 2020, le volume des fichiers non structurés progresse de 25% par an. En 2025, ils représentent déjà quatre cinquièmes du patrimoine informationnel d'une organisation française. Courriels archivés « pour mémoire », présentations estampillées « FINAL_V3 », images haute définition jamais compressées, logs applicatifs conservés « au cas où » : tout s'empile sans méthode ni horizon clair.

« Le COVID-19 a agi comme un accélérateur ; avec le télétravail chacun a stocké là où il le pouvait et le chaos s'est installé, constate Ludovic Kernalléguen, Architecte Logiciel, Spark Archives. Or l'illusion d'un disque illimité occulte un coût réel : stockage, réplication, sauvegarde, licences, énergie, audits de conformité ».

Face à cette marée, Spark Archives a développé ZERO VRAC, une plateforme SaaS qui cartographie, qualifie puis assainit les données dormantes sans interrompre l'activité. Notre promesse : transformer un problème que l'on croyait purement technique en avantage concurrentiel, réglementaire et environnemental.

À qui s'adresse ce contenu ?

Au DSI qui cherche de la capacité sans investir dans de nouveaux serveurs ou de nouvelles baies de stockage, au responsable conformité qui redoute la prochaine injonction de la CNIL, au responsable RSE qui inscrit la sobriété numérique dans son rapport CSRD, mais aussi à chaque métier qui perd du temps à chercher l'information juste… ou à la recréer.

Depuis vingt ans, Spark Archives accompagne les organisations françaises dans la maîtrise du cycle de vie documentaire. En 2025, le défi prioritaire n'est donc plus de produire davantage d'information ; mais de reprendre le contrôle sur l'avalanche de contenus non structurés – courriels, fichiers bureautiques, images, vidéos, copies de sauvegarde – qui parasitent les infrastructures et diluent la connaissance.

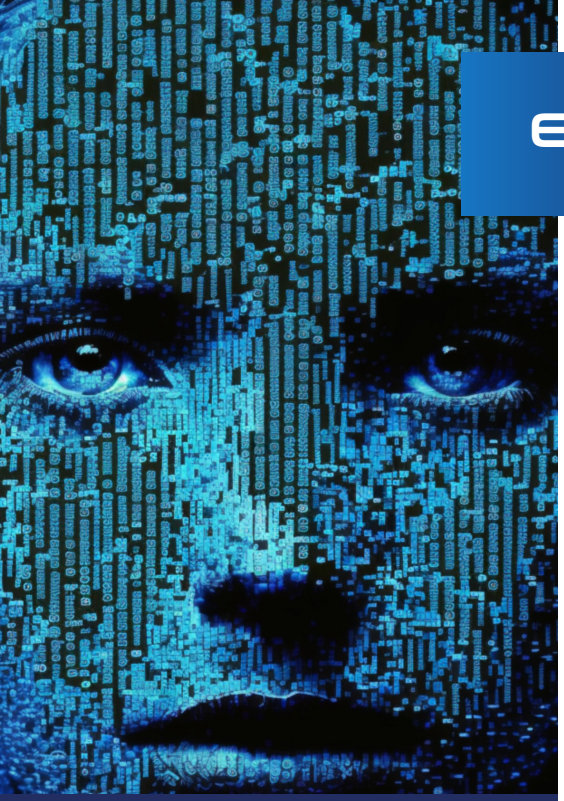
5 étapes clés

Ce livre blanc trace un itinéraire clair pour :
❶ **comprendre** l'ampleur du phénomène,
❷ **mesurer** ses impacts financiers, humains et climatiques,
❸ **démonter** les impasses des approches classiques,
❹ **découvrir** notre méthode éprouvée pour redevenir maître de son capital informationnel et éradiquer durablement le ROT (Redundant, Obsolete, Trivial) sans perturber les usages métiers,
❺ **quantifier** enfin le retour sur investissement et l'apport RSE.

Spark Archives entend être le porte-voix d'une gouvernance documentaire enfin tournée vers la valeur.

ZERO VRAC

Spark ARCHIVES
Une société de KLEE GROUP



Executive Summary

« Cartographier, qualifier, nettoyer, puis gouverner. ZERO VRAC en SaaS permet de tester sans projet lourd ; l'entreprise voit immédiatement la valeur » - Quentin Riac.

Les 4 bénéfices immédiats

❶ **Cash-flow :** baisse des coûts de stockage et de licence.

❷ **Agilité :** recherche documentaire accélérée, innovation facilitée.

❸ **Conformité :** réduction du risque de fuite ou de sanction.

❹ **RSE :** contribution tangible aux objectifs de neutralité carbone.

Les conseils d'administration ne disposent que de quelques minutes pour saisir l'essentiel. Cette page leur est dédiée. Voici de manière synthétique comment le vrac numérique, longtemps considéré comme un problème technique, se révèle un **levier de performance globale** lorsqu'il est traité de manière systémique.

Chaque salarié français génère aujourd'hui en moyenne 1 Go de données par jour1, dont 88% ne seront plus utilisées. Sur 100 To stockés, près de 60 sont déjà obsolètes ou redondants. Ces données saturent serveurs on-premise et souscriptions cloud, allongent les fenêtres de sauvegarde et multiplient les copies redondantes. Chaque mois de procrastination alourdit la facture énergétique et expose l'organisation à des risques de non-conformité (RGPD, NIS 2, DORA).

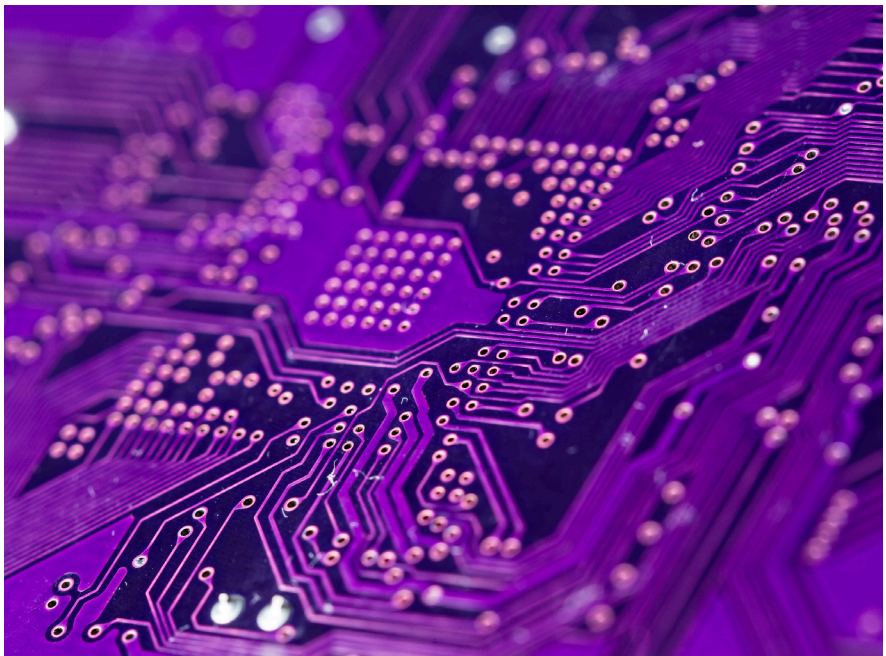
Assainir les volumes dormants sans déplacer les données

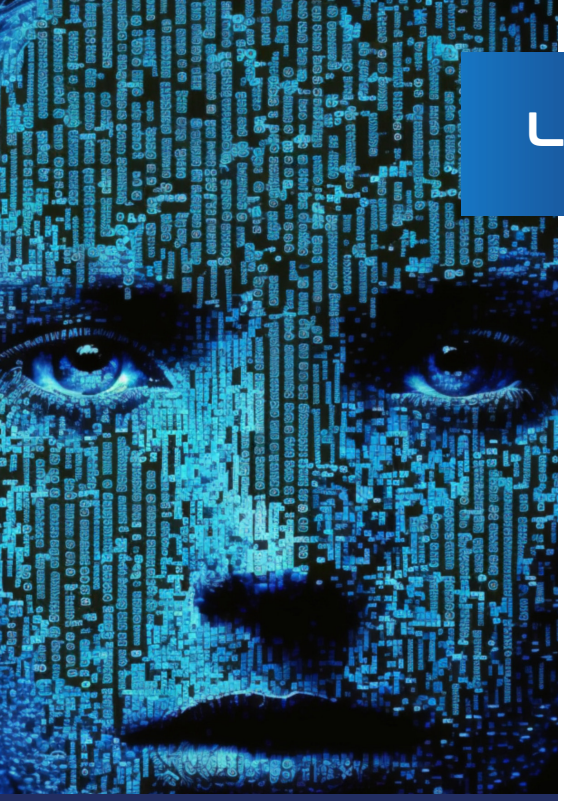
Spark Archives répond à cette problématique avec ZERO VRAC. La solution dresse une carte du patrimoine non structuré, évalue l'usage réel de chaque fichier, détecte les doublons, identifie les données personnelles et propose un plan d'assainissement que les métiers valident.

L'impact humain compte également. Quand la masse de données inutile recule, la recherche documentaire devient plus rapide et les équipes cessent de recréer l'existant.

L'argument environnemental achève de convaincre. 1 To de données stockées produit 25 kilogrammes de CO_2 par an. Supprimer 30 To revient à retirer 100 voitures de la circulation annuelle. Dans le reporting CSRD d'une entreprise, ce type d'indicateur pèse lourd.

Cette réduction de la masse informationnelle minimise également la surface d'attaque : moins de risques à gérer et moins de rançon potentielle à payer. Chaque euro économisé se double donc d'une tonne de carbone évitée et donc d'un risque réduit : la boucle vertueuse est complète.





« Quand le gaspillage devient visible, il devient négociable : une ligne budgétaire autrefois taboue devient un gisement d'économies » - Quentin Riac, Product Owner.

Les 4 bénéfices immédiats

- Backup complet > 10 h[2].
- Espaces SharePoint saturés à > 80 %[3].
- Facture cloud en hausse de +50% depuis deux exercices[4].
- Collaboration détournée vers des drives personnels (plus de 15 % des liens de partage émis via des outils grand-public)[5].

ZERO VRAC

Spark ARCHIVES
Une société de KLEE GROUP

Le défi dépasse la sphère IT

Pendant des années, l'informatique s'est concentrée sur la performance des serveurs, la bande passante réseau ou la disponibilité applicative. Or la **valeur informationnelle** se construit (ou se perd) dans les silos bureautiques et les espaces collaboratifs où s'entassent versions successives, documents orphelins et artefacts de projets clôturés.

Pression économique

Et ce n'est pas parce que le prix du téraoctet brut baisse, que l'affaire semble anodine. Car la réalité comptable, elle, raconte une tout autre histoire : réplications inter-sites, licences d'environnement collaboratif facturées au volume, sauvegardes incrémentales, PRA, énergie, etc. ; chaque gigaoctet inutile agit comme un passager clandestin qui consomme sans produire.
« Pourquoi dépenser des sous pour un patrimoine que votre organisation n'exploite plus ? »

Des enjeux humains

Cette pression financière s'accompagne d'un poids cognitif. De récentes études montrent que la recherche d'un document occupe désormais en moyenne 132 minutes par collaborateur et par semaine. Un chiffre qui grimpe à plus de 3 heures dans les services marketing, où la réutilisation de contenus est vitale.

La frustration engendrée par ces pertes de temps alimente un Shadow IT désormais endémique : échanges via WhatsApp, stockage temporaire dans des drives personnels, partage de « zip » WeTransfer jamais rapatriés. Autant de zones grises qui échappent à la DSI mais restent, juridiquement, sous la responsabilité de l'entreprise.
« Nettoyer avant de classer est souvent plus rentable que classer avant de nettoyer. » - Étude McKinsey 2024

Gouvernance et réglementations

Et la réglementation ne laisse plus de place à l'approximation. Le règlement DORA (Digital Operational Resilience Act), par exemple, exige, pour les établissements financiers, une cartographie précise des flux et une capacité de reprise en moins de deux heures.

La directive NIS 2 (sécurité des réseaux et des systèmes d'information), elle, étend l'obligation de résilience aux opérateurs d'importance vitale. D'ici janvier 2026, chaque fichier dormant devra être localisable, qualifié et rattaché à un cycle de vie. Sans gouvernance, le vrac numérique se transforme en passif réglementaire.
« Même avec une GED, le risque subsiste ; le collaborateur contourne l'outil dès qu'il ne trouve pas son compte »

Au-delà de la conformité, la responsabilité sociétale amplifie l'urgence. Le scope 2 des bilans carbone intègre déjà l'électricité consommée par les datacenters. Demain, le scope 3 comptabilisera la fabrication des supports et la fin de vie des serveurs. L'inaction revient donc à subventionner un double impact négatif : financier et environnemental.

Anatomie d'un chaos informationnel



« Identifier les doublons parmi 10 millions de fichiers exigerait plus de 3 ans de lecture humaine à raison de 2 secondes par fichier : l'opération est rigoureusement impossible sans automatisation ».

Identifier les mécanismes de prolifération

❶ Versioning sans gouvernance : Project-final-v3-def-final2.docx.

❷ Shadow IT : échanges hors périmètre via wetransfer, Slack, etc.

❸ Parachutage de données après fusion-acquisition.

❹ Absence de politique de rétention : « On garde, au cas où ».

Chaque partage réseau est un mille-feuille. La strate 2014 contient encore les PST issus de la messagerie Lotus Notes, la strate 2017 réunit des captures d'écran HD déposées pour un hackathon, la strate 2020 voit fleurir des fichiers Teams exportés en urgence pendant le confinement. La cartographie d'un dépôt type de 15 To illustre cette réalité :

- 5 To d'archives PST/OST[1] issues de messageries historiques (notamment Outlook),
- 4 To de documents Office (dont 42 % de duplications strictes ou proches),
- 3 To d'images et de vidéos marketing, résolution et usage inconnus,
- 2 To de logs applicatifs zippés,
- 1 To de sauvegardes « temporairement » stockées… depuis trois ans.

Les causes premières du chaos sont connues, mais rarement traitées : versioning sans règle, mobilité non encadrée, projets de fusion-acquisition sans plan d'intégration documentaire, vision comptable du stockage (« moins cher chaque année »). Tant que l'on ne renchérit pas artificiellement le coût du gigaoctet, le réflexe « On garde, au cas où » domine.

À ce stade d'enchevêtrement, l'examen humain n'est plus envisageable. ZERO VRAC est là pour repérer les copies strictes, et sera bientôt capable d'appliquer un algorithme de similarité pour isoler les « faux jumeaux », ces présentations rebaptisées final_final_v2.pptx dont seules trois slides diffèrent.

Agir avec ZERO VRAC

Il est pourtant possible d'engager une qualification structurée mais volontairement prudente du patrimoine analysé. À ce stade, ZERO VRAC s'appuie uniquement sur des éléments objectivables et immédiatement accessibles : typologies de fichiers, volumes, formats, arborescences, dates de création et de modification, ainsi que l'analyse des doublons stricts et sémantiques. Cette première lecture suffit à mettre en évidence des gisements de données clairement redondantes ou obsolètes, sans extrapoler des usages qui ne peuvent être techniquement garantis.

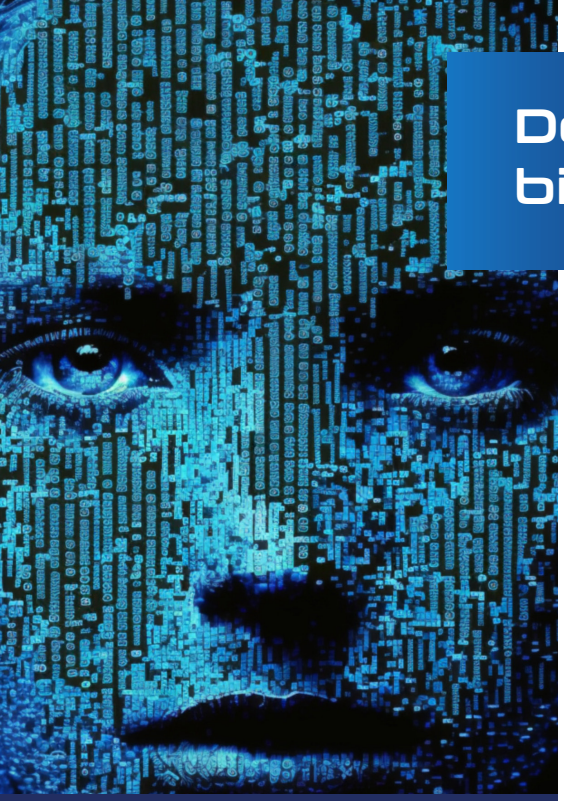
Certaines dimensions, fréquemment associées aux projets de gouvernance documentaire, sont délibérément exclues de l'analyse actuelle afin d'éviter toute approximation. La date de dernière lecture, la fréquence de partage ou le profil du producteur ne sont pas exploités tant que les connecteurs applicatifs correspondants ne sont pas en place. De même, la détection automatique de données personnelles fait partie de la roadmap produit, mais n'est pas activée dans la version actuelle.

Ce positionnement assumé permet de fournir une aide à la décision fiable, fondée sur des critères mesurables et vérifiables. Les enrichissements à venir (accès aux journaux d'usage, contextualisation utilisateur, identification fine des PII) viendront compléter progressivement cette cartographie, sans remettre en cause les analyses déjà produites ni la valeur immédiate délivrée.

[1] Personal Storage Table/Offline Storage Table

ZERO VRAC

Spark ARCHIVES
Une société de Kleegroup



Des impacts cachés, mais bien réels

« Le plus grand gain n'est pas le gigaoctet supprimé, mais la décision qu'on prend enfin sur la valeur de l'information » - Quentin Riac.

Quadruple impact :

- 2h/sem/salarié perdues à chercher des fichiers.
- 900 €/To/an coûts complets.
- 65 % des ransomwares exploitent les partages de fichiers.
- 25 kg CO_2/To/an émis.

La colonne « stockage » du budget IT cache une forêt d'externalités. À 900€/an le téraoctet inutilisé, une ETI approche le demi-million lorsqu'elle dépasse 500 To. Et encore, ce n'est qu'un début. Le cabinet Forrester calcule qu'un salarié du tertiaire reproduit en moyenne 8 % de son travail faute de retrouver la version adéquate. Sur un effectif d'un millier de collaborateurs facturés 55 k€ par an, la perte dépasse 4,4 M€ : une ligne intangible car considérée comme « temps de recherche normal ».

La surface d'exposition cyber suit une logique mathématique : plus le pool de données et leur partage sont volumineux, plus la quantité de points d'entrée augmente. Les rapports Cisco Talos 2025 démontrent que les ransomwares ciblent prioritairement les volumes peu gouvernés, gage de récupération lente donc rançon élevée. ZERO VRAC agit ici de manière indirecte mais déterminante. En réduisant en amont les volumes inutiles (doublons, archives dormantes, données obsolètes, etc.) l'organisation diminue mécaniquement la quantité de données à la fois exposées au chiffrement et à la restauration.

En cas de ransomware, assainir le périmètre documentaire réduit mécaniquement l'ampleur de la crise : restaurer 50 To sera toujours plus rapide, moins coûteux et moins risqué que restaurer 100 To. Moins il y a de données inutiles à sauvegarder, chiffrer ou restaurer, plus le retour à la normale est maîtrisable. ZERO VRAC ne bloque pas l'attaque, mais il réduit son rayon d'impact opérationnel, transformant un arrêt prolongé en incident maîtrisé.

Quant au volet environnemental, il n'est plus cosmétique. Avec 25 kilogrammes de CO_2 par To et par an, l'assainissement de 50 To représente l'équivalent annuel de 170 000 kilomètres parcourus en voiture citadine. Les rapports CSRD intégreront ce chiffre noir sur blanc. Anticiper, c'est raconter une histoire positive au marché et aux talents.

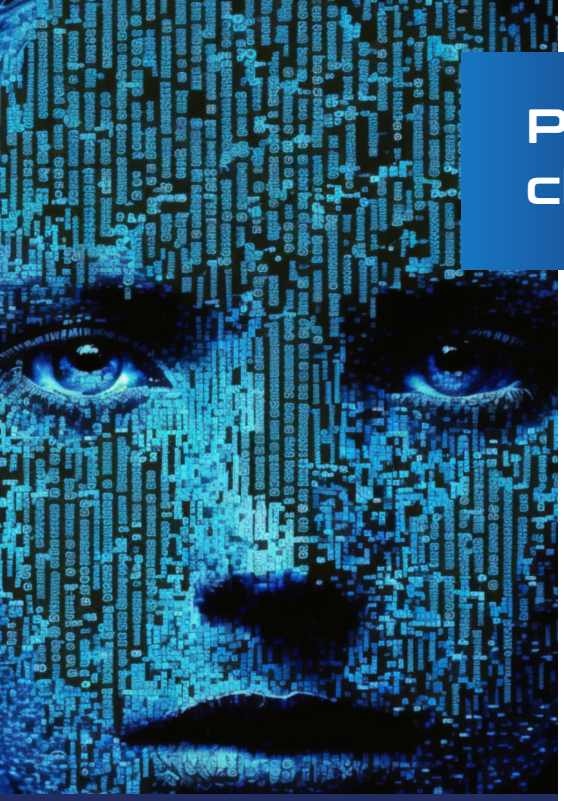
Sans oublier la question de la conformité, et des angles morts que cette gestion anarchique de l'information créent sur les données personnelles ou sensibles.



ZERO VRAC

Spark ARCHIVES

Une société de KLEE GROUP



Pourquoi les remèdes classiques échouent ?

« Le classement manuel se heurte à la loi de Pareto : 20 % des collaborateurs tiennent 80 % des règles, les autres abandonnent devant la lourdeur.»

Les clichés battus en brèche :

- “La GED suffit”
- “Le cloud est infini”
- “On nettoiera plus tard”.

La tentation de la solution unique et centralisatrice est forte : GED/ECM toute-puissante, script de purge mensuel ou grand déménagement vers un cloud « élastique ». Si les plateformes de GED/ECM gèrent à merveille les flux entrants, elles n’abordent jamais le stock hérité, faute de pouvoir ingérer et traiter 10 ou 50 To détériorés.

Elles exigent un effort de classement manuel et n’offrent pas de moteur de détection des doublons ou d’analyse d’obsolescence. Depuis toujours les entreprises conservent “au cas où”. Les migrations de masse, elles, créent des copies transitoires qu’on oublie d’effacer, doublant le volume et la facture.

Classement manuel

La tentation de « demander aux utilisateurs de ranger », se heurte à la réalité : chaque employé devrait passer 15 % de son temps à décrire des métadonnées exhaustives pour atteindre un niveau satisfaisant de gouvernance. Aucune organisation ne peut accepter ce coût caché. Il est possible d’ajouter des métadonnées manuellement, mais personne ne le fait, parce que c’est rébarbatif et beaucoup trop long.

Politiques de rétention génériques

Et si la règle « Effacer après 7 ans » était la solution ? Cela semble simple, mais la valeur métier d’un document n’est jamais parfaitement corrélée à son âge. Résultat : la règle est suspendue par peur de perdre un contenu clé, ou bien appliquée à la lettre et aboutit à la suppression de données critiques. Rappelons qu’un plan stratégique peut courir sur 10 ans, une instruction judiciaire peut survenir 15 ans après les faits, une base R&D reste pertinente 20 ans, etc. Sans lecture contextuelle, la règle unique détruit parfois l’actif le plus précieux.

Quant aux scripts maison, ils montrent vite leurs limites : absence de moteur sémantique, manque de journalisation, impossibilité de gérer les exceptions métier, obsolescence du code lorsque l’admin change de poste.

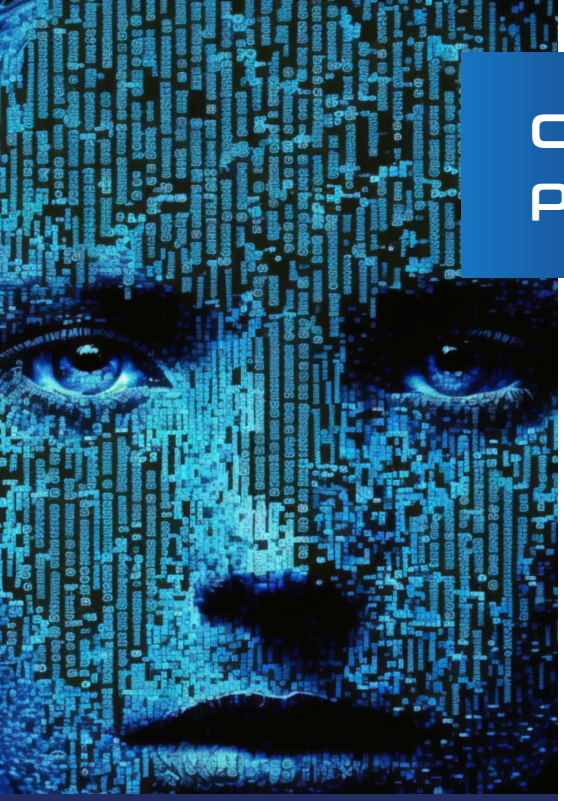
ZERO VRAC évite ces ornières en dissociant détection, gouvernance et action. La solution identifie et propose, les métiers disposent. L’outil n’impose pas un classement ; il fournit un miroir objectif dans lequel chaque direction se regarde et décide.



ZERO VRAC

Spark ARCHIVES

Une société de KLEE GROUP



Complexité hybride et double peine

« Le cloud n'est pas une déchetterie gratuite ; chaque giga inerte déposé devient un loyer perpétuel ».

L'addition cachée :

- Jusqu'à 0,30 € sur chaque euro cloud peut partir dans le stockage de doublons, fichiers périmés, documents sans valeur.
- 48 % des fichiers migrés ne seront jamais rouverts

L'environnement informationnel 2025 ressemble à une mosaïque mouvante : Microsoft 365, Google Workspace, Salesforce, Slack, Confluence, NetApp on-prem, buckets S3, blob Azure, etc. Chaque brique possède son propre mécanisme d'indexation, son modèle de permission, sa logique de version, son moteur de recherche et son vocabulaire. L'utilisateur bascule de l'un à l'autre en toute transparence, mais la gouvernance y perd sa cohérence.

Face à l'inertie des processus internes, les équipes ont tôt fait d'adopter des solutions externes (Discord, Telegram, Google Drive, Box, Dropbox, etc.). Ces « dark repositories » ne sont pas référencés et échappent au contrôle, mais restent dans le périmètre de responsabilité RGPD. En cas de fuite, impossible d'arguer de l'ignorance : la responsabilité demeure.

Cloud ≠ sortie de secours

Le cloud, souvent présenté comme la panacée, déplace plus qu'il ne résout. Les fournisseurs de services cloud facturent d'abord le stockage, puis des frais d'egress lorsque l'on souhaite récupérer ou migrer. Chaque téraoctet de ROT transféré, c'est un double coût immédiat, puis un triple lorsqu'il faudra le nettoyer. Quentin Riac résume : « Transporter du vrac, c'est déménager des cartons jamais ouverts ». Migrer vers le cloud revient donc à déplacer le problème.

ZERO VRAC adresse cette dispersion en peaufinant actuellement son intégration native aux outils existants, de Sharepoint à la GED en passant par le cloud.



ZERO VRAC

Spark ARCHIVES
Une société de KLEE GROUP

ZERO VRAC : la solution qui change



« ZERO VRAC ne remplace pas la décision métier : il la rend possible, rapide et documentée. »

Ce qu'il faut retenir

ZERO VRAC permet de créer des périmètres d'analyse ciblés, d'y intégrer des documents depuis plusieurs sources, d'extraire contenus et métadonnées, d'identifier doublons et similarités, d'enrichir les documents par étiquettes, règles, mots-clés ou référentiels métiers, puis de préparer leur export, leur mise à jour, leur suppression contrôlée ou leur versement vers une application cible.

La promesse de ZERO VRAC tient en une phrase : transformer un stock documentaire indifférencié en patrimoine informationnel lisible, qualifié et actionnable. L'objectif n'est pas de remplacer les outils existants, ni d'imposer une nouvelle organisation documentaire, mais d'ouvrir un espace d'analyse, de tri et d'arbitrage dans lequel l'organisation peut enfin voir ce qu'elle possède, comprendre ce qui l'encombre et décider ce qui doit être conservé, archivé, enrichi, déplacé ou supprimé.

ZERO VRAC part d'un constat simple : le vrac numérique ne se traite pas uniquement avec du stockage, ni avec une règle de purge générique. Il faut d'abord rendre le désordre observable. La solution permet donc de créer autant de périmètres de travail que nécessaire, appelés "vracs", correspondant à un dossier, un fonds documentaire, une direction métier, un espace SharePoint, un ensemble de fichiers bureautiques ou un stock hérité à remettre sous contrôle.

Une alimentation progressive des contenus

Les documents peuvent être intégrés dans ZERO VRAC selon plusieurs modalités. Le dépôt peut d'abord se faire de manière simple, par téléversement manuel de documents ou de dossiers. Il peut également passer par un client sécurisé permettant de déposer les fichiers dans un bucket S3, après authentification. Spark Archives travaille aussi à l'automatisation de l'alimentation depuis des applications distantes, notamment SharePoint, grâce à une orchestration de type Apache NiFi. Cette approche permet de récupérer les contenus et leurs métadonnées associées, puis de les faire entrer dans le pipeline d'analyse de ZERO VRAC.

Cette logique est essentielle. La solution construit un environnement de travail maîtrisé où les documents sont analysés selon des règles homogènes, quel que soit leur point d'origine. Elle permet ainsi de comparer, qualifier et prioriser les contenus sans dépendre des limites techniques propres à chaque application source.

Du dépôt au diagnostic

Une fois les documents intégrés, ZERO VRAC les fait passer dans un pipeline d'analyse. Celui-ci extrait les métadonnées disponibles, récupère le contenu exploitable, génère une vignette et produit une prévisualisation PDF des premières pages. L'utilisateur peut ainsi vérifier rapidement la nature d'un document sans sortir de l'écran de travail ni télécharger systématiquement les fichiers.

La solution détecte ensuite les doublons stricts, c'est-à-dire les fichiers possédant la même empreinte, mais aussi les documents similaires. Cette deuxième capacité est déterminante pour traiter les faux jumeaux documentaires : versions successives, fichiers renommés, présentations légèrement modifiées, documents "final", "final_v2" ou "définitif" dont seules quelques parties diffèrent. ZERO VRAC peut ainsi faire apparaître des proximités de contenu et aider les métiers à choisir quelle version conserver.

ZERO VRAC

Spark ARCHIVES

Une société de KLEE GROUP



Le diagnostic ne se limite pas aux doublons. Les documents peuvent être analysés et filtrés par familles de formats, poids, langue, mots-clés, métadonnées, utilisateurs, étiquettes ou classeurs. L'objectif étant de donner à l'utilisateur une vision macro exploitable : où sont les gros volumes, quels formats dominent, quels contenus se répètent, quels ensembles doivent être traités en priorité.

Une qualification métier, pas seulement technique

La force de ZERO VRAC tient aussi à sa capacité à rapprocher l'analyse documentaire de la réalité métier. Les utilisateurs peuvent ainsi créer des étiquettes, les appliquer à un document ou à un ensemble de documents, les combiner avec d'autres critères, puis filtrer les résultats. Dans une logique collaborative, les étiquettes peuvent être enrichies, fusionnées ou rationalisées afin d'éviter la prolifération de tags équivalents.

La notion de "classeur" permet ensuite de préparer le sort final des documents : conserver, archiver, détruire, revoir, traiter selon une durée spécifique ou orienter vers un autre circuit. Un document n'a vocation à appartenir qu'à un seul classeur, ce qui évite l'ambiguïté au moment de décider. ZERO VRAC peut également affecter des documents à des utilisateurs ou à des responsables, afin de répartir le travail de qualification entre équipes métiers, DSI, record managers ou archivistes.

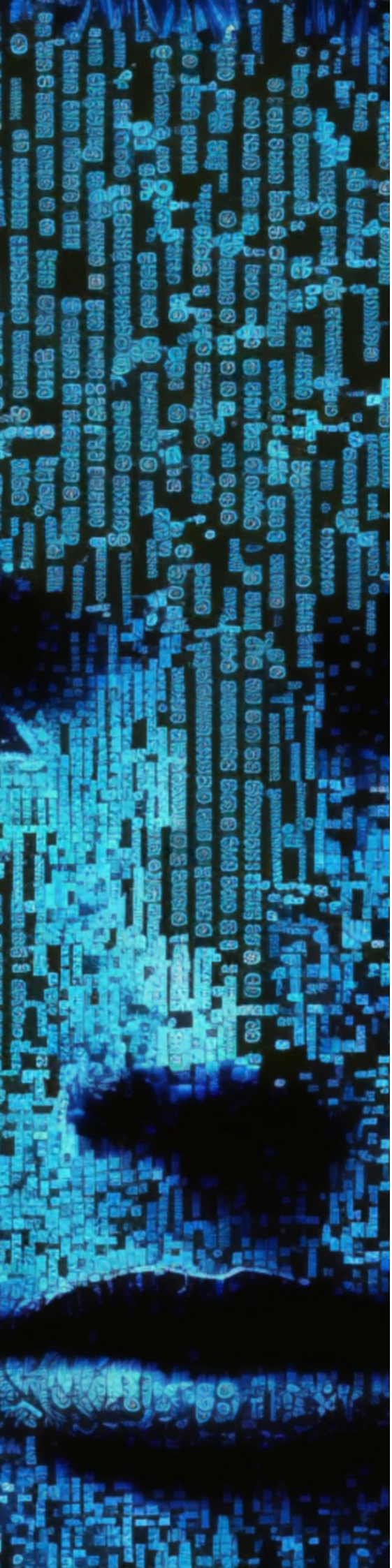
La solution intègre par ailleurs une logique de référentiels. Un thésaurus RDF/SKOS peut être importé afin d'analyser les contenus à partir d'un vocabulaire contrôlé (termes préférentiels, synonymes, termes alternatifs, notions reliées). Dans un secteur comme la pharmacie, par exemple, ce mécanisme permet d'identifier rapidement les documents associés à une molécule, une famille thérapeutique ou un sujet réglementaire sans dépendre uniquement d'une recherche plein texte classique.

Enrichir pour mieux décider

ZERO VRAC ne se contente pas de lire les documents ; il prépare aussi leur gouvernance. Un moteur de règles permet de créer de nouvelles métadonnées à partir de critères définis : type de document, étiquette, format, contenu détecté, règle métier, etc. Ces métadonnées enrichies peuvent ensuite servir au tri, à l'export, au reclassement ou à la mise à jour des systèmes cibles.

Des traitements par IA peuvent également être appliqués à certains contenus non textuels, notamment des images et d'autres cas d'suages sont en préparation. Lorsqu'un fichier ne contient pas de texte exploitable, l'analyse peut produire une description, une catégorisation et des mots-clés. Cela ouvre la voie à une meilleure recherche dans des fonds où les images, scans ou fichiers visuels restent souvent invisibles pour les moteurs classiques.

La recherche s'appuie enfin sur le contenu extrait et les métadonnées disponibles. Elle permet de croiser les critères, d'isoler des corpus cohérents, d'identifier les documents à traiter et de bâtir des lots d'action fiables.



Une solution d'arbitrage avant d'être une solution d'exécution

ZERO VRAC ne doit pas être présenté comme un bouton magique de suppression. Sa valeur réside dans sa capacité à fournir une base d'arbitrage documentée. La solution aide à identifier, organiser et catégoriser les documents pour permettre aux métiers et à l'IT de décider. Dans certains cas, une DSI cherchera d'abord à supprimer les doublons pour réduire fortement son stock. Dans d'autres, un service conformité voudra enrichir les métadonnées, préparer un versement en SAE ou réorganiser un fonds documentaire.

L'export devient alors une étape clé. ZERO VRAC permet de générer des exports locaux, de préparer la suppression de documents identifiés comme doublons, de mettre à jour des métadonnées dans l'application d'origine lorsque le connecteur le permet, ou d'envoyer les documents et leurs métadonnées vers une autre application, comme un SAE ou un autre espace SharePoint.

ZERO VRAC rend le vrac visible, qualifie les contenus et prépare l'action. L'entreprise garde la main sur les décisions.

Alimenter ZERO VRAC

Dépôt manuel, client sécurisé/S3, alimentation connectée depuis les applications sources.

↓

Analyser le corpus

Extraction de métadonnées, prévisualisation, indexation, doublons stricts, similarités.

↓

Qualifier et arbitrer

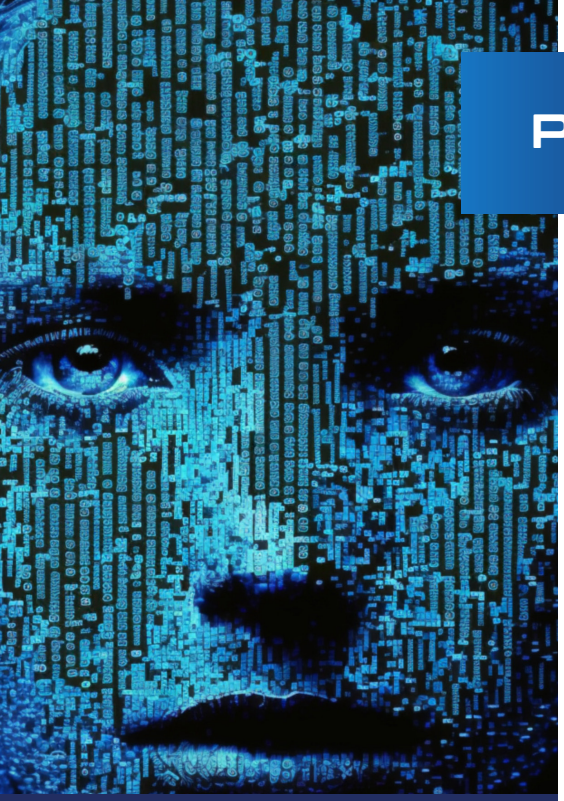
Étiquettes, mots-clés, référentiels SKOS, règles d'enrichissement, classeurs, affectation utilisateurs.

↓

Agir et gouverner

Export local, mise à jour de métadonnées, préparation de suppression contrôlée, versement vers SAE ou application cible.

Spark Archives travaille également à une fiche de vie existante et opérationnelle permettant de suivre l'avancement du traitement d'un vrac : part des documents analysés, enrichis, qualifiés, affectés ou prêts à l'export. Cette vision de progression est importante pour piloter les campagnes d'assainissement dans la durée et objectiver les gains obtenus.



« Nous pensions devoir acheter de nouveaux nœuds de stockage. En réduisant le vrac numérique, nous avons finalement repoussé cet investissement. »

Pourquoi ZERO VRAC ?

- UX adoptée en < 30 min ;
- intégration native Microsoft 365, GED & cloud ;
- ROI moyen < 12 mois.

ZERO VRAC

Spark ARCHIVES
Une société de KLEE GROUP

Pourquoi choisir ZERO VRAC ?

Choisir un outil de gouvernance documentaire engage l'entreprise pour des années ; le moindre irritant freine l'adoption et ruine le retour sur investissement. ZERO VRAC a été conçu pour lever ces blocages dès la première prise en main.

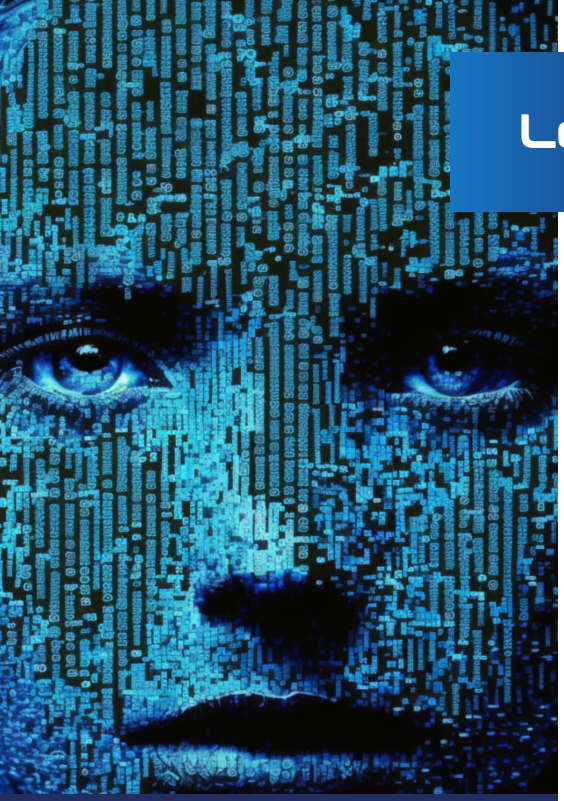
L'expérience utilisateur s'aligne sur les gestes quotidiens : un simple tableau de bord web, aucune modalité déroutante, aucune fenêtre intrusive. Les métiers visualisent leur patrimoine informationnel à travers des indicateurs simples et immédiatement compréhensibles. Une jauge de remplissage permet notamment de suivre, en temps réel, le taux d'occupation du tenant ou de l'espace analysé, et de matérialiser la pression exercée par le vrac numérique sur les capacités disponibles. Cette visualisation rend le problème concret : l'espace se remplit, les marges se réduisent, les arbitrages deviennent nécessaires. L'interface ne prétend pas automatiser magiquement les décisions, mais elle fournit une lecture claire et partagée de la situation.

Mais la simplicité d'usage ne saurait suffire sans une différenciation technologique crédible et assumée. ZERO VRAC concentre aujourd'hui ses efforts sur des mécanismes de similarité documentaire permettant d'identifier les doublons stricts et les documents très proches, afin de rendre visibles les redondances les plus coûteuses. L'objectif est volontairement ciblé : fournir des indicateurs fiables sur la volumétrie, la répétition et la structure des contenus, sans surinterprétation.

Cette approche pragmatique permet d'obtenir rapidement des résultats exploitables et de poser les bases d'une gouvernance documentaire progressive, fondée sur des faits mesurables plutôt que sur des promesses technologiques prématurées.



Le vrac numérique n'attend pas



« Le vrac numérique est une dette technique devenue dette carbone. S'en libérer, ce n'est pas seulement réduire les volumes stockés, c'est aussi accélérer la recherche d'information, simplifier les audits, raccourcir les restaurations après incident et alléger durablement l'infrastructure. Autrement dit, gagner en vitesse autant qu'en sobriété. »

En résumé

- Vrac numérique = risque + coût + CO_2.
- ZERO VRAC = ROI + conformité + sobriété.
- Agissez dès aujourd'hui, sécurisez demain.

L'explosion annoncée des contenus générés par l'IA va transformer le vrac numérique en tsunami. Sans oublier que les réglementations DORA et NIS 2 imposent désormais une traçabilité exhaustive des actifs d'information et des chaînes d'approvisionnement numériques.

Sans discipline, chaque prompt créera donc son document, chaque révision créera sa version, chaque sauvegarde son clone, chaque IoT son lot de logs redondants. Les organisations capables d'assainir automatiquement resteront les seules à exploiter sereinement ces nouvelles sources de valeur.

Spark Archives propose une réponse concrète : ZERO VRAC transforme votre passif documentaire en un actif gouverné, sécurisé et sobre. L'investissement est financé par les économies qu'il génère, et ses bénéfices – financiers, opérationnels, environnementaux – sont immédiats.

ZERO VRAC n'est ni un énième outil de dé-duplication ni un grand soir documentaire ; c'est une démarche progressive, mesurable, pilotée par la valeur. Le financement provient des économies réalisées : le bénéfice se cristallise dans la productivité humaine, la preuve carbone et la résilience cyber.

Avant de refermer ce contenu, posez-vous questions :

- Savez-vous où dorment vos données ?
- Pouvez-vous facilement les localiser et les qualifier ?
- Saurez-vous prouver leur traçabilité en 2026 ?
- Êtes-vous prêt à payer 25 kg CO_2 par téraoctet et par an pour du contenu inutile ?

Si une seule réponse est hésitante, le moment est venu d'agir.

Ensemble, transformons la dette technique et la dette carbone en un atout business durable.

Passez à l'action

1. Scannez le QR code pour programmer une démo personnalisée.
2. Demandez votre diagnostic flash.

Vous avez maintenant les clés pour mettre fin au vrac numérique. La seule question qui subsiste : **qu'attendez-vous ?**

Spark Archives au service de la gestion documentaire depuis plus de 15 ans.