Microsoft 365, IA et données non structurées : protéger vos données d'affaires

Livre électronique

WeActis
par Mondata

Sécuriser les données non structurées dans un milieu de travail axé sur l'IA

Notre façon de travailler a changé. Les organisations dépendent de plus en plus d'une solide collaboration, des outils infonuagiques et de l'automatisation alimentée par l'IA pour demeurer concurrentielles. Microsoft 365 est devenu la pierre angulaire des communications d'affaires, générant et hébergeant de **vastes volumes de données non structurées** – comme des courriels, des documents, des clavardages et plus encore. Bien que ces progrès améliorent l'efficacité, ils posent des défis considérables en matière de sécurité, de conformité et de gouvernance des données.

L'intelligence artificielle accélère les occasions et les risques, **ce qui exacerbe les vulnérabilités existantes** en matière d'accès aux données, de partage et de contrôle. Sans une stratégie de préparation claire, les organisations risquent d'exposer des renseignements sensibles, d'agrandir leur surface d'attaque et de devenir la cible de menaces, comme les rançongiciels et les fuites de données.

Ce livre électronique offre une **feuille de route** pour sécuriser les données non structurées dans un milieu de travail axé sur l'IA, couvrant :

- **L'intersection de l'intelligence artificielle et des données non structurées – opportunités et menaces.**
- **Les principaux risques liés à une collaboration avec Microsoft 365 et le rôle de l'IA pour les exacerber.**
- **Stratégies éprouvées pour améliorer le contrôle de l'accès, l'hygiène des données et la conformité.**
- **Mesures concrètes pour renforcer la cyberrésilience sans perturber les opérations de l'entreprise.**

Êtes-vous prêt à prendre le contrôle de vos données non structurées ? Allons-y.

Données non structurées, intelligence artificielle et défis en matière de sécurité

L'intelligence artificielle (IA) a rapidement transformé la façon dont les entreprises fonctionnent, et les organisations sont désireuses de tirer parti de son potentiel pour améliorer l'efficacité et demeurer concurrentielles. Cependant, l'intégration poussée de l'intelligence artificielle dans des outils de collaboration comme Microsoft 365 présente de nouveaux défis en matière de sécurité et de conformité, surtout en ce qui concerne les données non structurées.

Chaque jour, les entreprises génèrent **d'énormes quantités de données non structurées** – documents, courriels, messages de clavardage, feuilles de calcul, etc.

Bien que les plateformes de collaboration améliorent la productivité, elles exposent les organisations à des risques, comme des atteintes à la protection des données, des violations de la conformité et des cybermenaces, comme les rançongiciels. L'intelligence artificielle amplifie davantage ces risques en automatisant le traitement des données, parfois sans visibilité ou contrôle clair sur les renseignements sensibles.

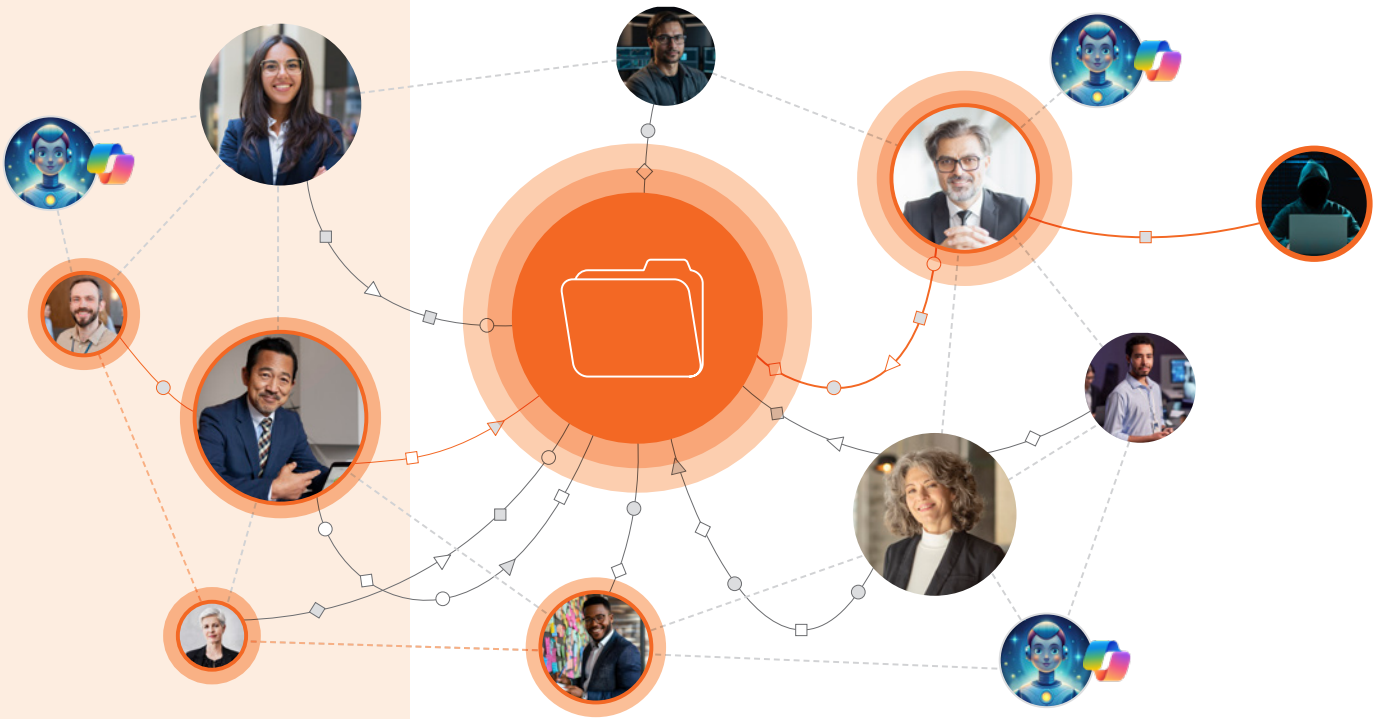
À mesure que l'IA transforme le milieu de travail numérique, les organisations doivent adapter proactivement leurs stratégies en matière de sécurité, de risque et de conformité.

L’ascension des données non structurées

Quels trésors se cachent dans ces environnements collaboratifs ?

La réponse est simple, ce sont les données non structurées.

Par le passé, nous avons considérablement amélioré la sécurité des **données structurées.** Ce type de données se trouve généralement dans les applications et les systèmes. Elles sont souvent planifiées et organisées, ce qui en rend le contenu un peu plus prévisible, bien qu’il puisse encore y avoir des surprises. Bien qu’il s’agisse d’un défi permanent, on considère désormais les enjeux liés aux données non structurées comme étant plus importants.



Les données non structurées sont disponibles dans différents formats de fichier.

Elles peuvent comprendre des images, des documents Word, des fichiers Excel, des présentations, de l'audio et plus encore. Par conséquent, c'est toujours différent et, pourrait-on dire, imprévisible. L'adoption d'environnements collaboratifs et plusieurs autres facteurs ont fait grimper la quantité de données non structurées à un taux de croissance annuel renversant de 55 % à 65 %. Certains emploient l'expression « **explosion** » pour décrire la situation actuelle.

Il est clair que les données sont devenues la nouvelle monnaie d'échange et que de nombreuses innovations visent à générer plus de données, même s'il y a souvent un dédoublement considérable. De nombreuses initiatives de gouvernance, en particulier dans le secteur public, ont établi une certaine structure pour catégoriser ces données afin d'assurer une protection adéquate. Cet effort vise à guider les organisations et à gérer efficacement les risques associés à ces données.

Les données non structurées augmentent au taux annuel de 55 % à 65 % et représentent plus de 80 % de toutes les données d'entreprise

Congruity 360 - 2023

Rançongiciel – une industrie d'un milliard de dollars

La croissance des données non structurées a également causé des maux de tête dans le domaine de la cybersécurité et d'autres industries. Les criminels ont reconnu et saisi l'occasion que présentent les données non structurées et la dépendance des organisations à leur égard pour perturber les opérations et générer des revenus, que ce soit en les rendant inaccessibles à l'organisation (par le chiffrement) ou en les exposant publiquement (au moyen de fuites), forçant ainsi l'entreprise à verser une rançon.

Nous avons observé une croissance importante du marché des systèmes et des applications qui aident les organisations à comprendre les données non structurées, conformément au principe de cybersécurité selon lequel il faut savoir ce qui doit être sécurisé. Ces systèmes peuvent examiner les magasins de données d'une organisation, produire des rapports sur les types de données repérées et, dans certains cas, même étiqueter les fichiers et appliquer certaines mesures de contrôle.

Les entreprises ont besoin de stratégies solides pour gérer la quantité croissante de données potentiellement sensibles; des solutions de cybersécurité conçues pour cette situation devraient être mises en œuvre rapidement afin d'assurer une utilisation optimale et de prévenir les incidents coûteux liés à la cybersécurité et à la protection des renseignements personnels.

Données non structurées et IA générative

Un système d'intelligence artificielle se définit comme une construction mathématique capable de prédire ou de créer du contenu à partir des informations initiales. Il existe de nombreux liens directs entre les données non structurées et l'IA générative.

1 Le premier point est assez fondamental : **la plupart des modèles d'IA générative existent parce qu'ils sont formés sur les données non structurées.** Ces modèles sont fondés sur de grands modèles de langage (GML, nommé LLM en anglais).

Par exemple, un GML (grand modèle de langage) tel que ChatGPT consiste en un réseau de neurones conçu pour comprendre et générer du texte de manière cohérente. Il est formé sur de grandes quantités de texte pour apprendre les relations entre les mots et les phrases.

Les paramètres sont des variables internes du modèle qui sont ajustées pendant la formation (ou l'apprentissage) afin d'optimiser les prévisions. Les GML possèdent des milliards de paramètres qui influencent la façon dont le modèle interprète une entrée de texte et génère une sortie. Ces paramètres sont ajustés pendant l'apprentissage du modèle afin de réduire les erreurs de prédiction liées aux données de formation.

2 Deuxièmement, **la plus grande force de l'IA générative est d'aider à générer des données non structurées.**

Comme nous l'avons mentionné, les GML sont conçus pour prédire le résultat suivant, ce qui les rend très efficaces pour générer du contenu à partir de données non structurées lorsqu'on leur demande. Pensez à votre agent de clavardage, vos assistants virtuels ou d'autres applications qui vous aideront à rédiger du contenu. Cela contribue de manière significative à la croissance évoquée précédemment, car il ne s'agit plus seulement d'utilisateurs, mais aussi de systèmes capables de créer du contenu. De plus, ces systèmes ne dorment pas et ne mangent pas, ce qui leur permet de fonctionner en continu 24 heures par jour, 7 jours par semaine.

3 Troisièmement, **l'IA générative permet de naviguer habilement dans un labyrinthe de données non structurées, que ce soit à des fins légitimes ou potentiellement malveillantes.**

Lorsque vous « interagissez » avec un assistant en intelligence artificielle dans un environnement de collaboration, il fait d'abord l'objet d'un processus d'« ancrage ». Ce processus harmonise les réponses d'un modèle avec des données exactes et réelles afin d'assurer la pertinence, la fiabilité et l'uniformité. Il recueille des renseignements de diverses sources propres à l'organisation. Selon sa configuration, les sources et les permissions d'accès peuvent varier; toutefois, dans un environnement Microsoft 365, cela pourrait comprendre « toutes » données auxquelles l'utilisateur final peut accéder, comme des partages d'autres utilisateurs, des accès antérieurs et plus encore.

Lorsqu'un utilisateur ayant un accès non autorisé invite l'assistant Copilot de Microsoft 365 à poser une question, **la réponse pourrait facilement révéler des renseignements de nature délicate que ce même utilisateur ne connaissait pas et auxquels il n'avait pas accès.**

La collaboration a un coût

Les risques croissants liés aux données non structurées

Faute de connaissances des utilisateurs en hygiène des données, les TI héritent d'un désordre qui s'accumule, sans le contexte d'affaires pour y remédier.

Révision des accès

Étant donné qu'il y a beaucoup plus de données qu'auparavant et qu'elles sont en grande partie « non structurées », il est plus important que jamais de maintenir une révision régulière, à jour et minimale des accès. Ce processus doit avoir lieu régulièrement pour chaque bibliothèque où les données sont stockées et, idéalement, **être mené par des utilisateurs qui correspondent étroitement aux objectifs opérationnels de ces bibliothèques.**

Partages

Les partages sont des accès **« discrétionnaires »** créés pour les utilisateurs qui ne font pas partie du public habituel indiqué dans les révisions des accès standard dont il a été question précédemment. Ils sont excellents pour la productivité, mais devraient être **temporaires** et **limités**. Nous créons constamment des partages; nous devrions les examiner et les révoquer régulièrement pour ne conserver que ceux qui sont nécessaires.

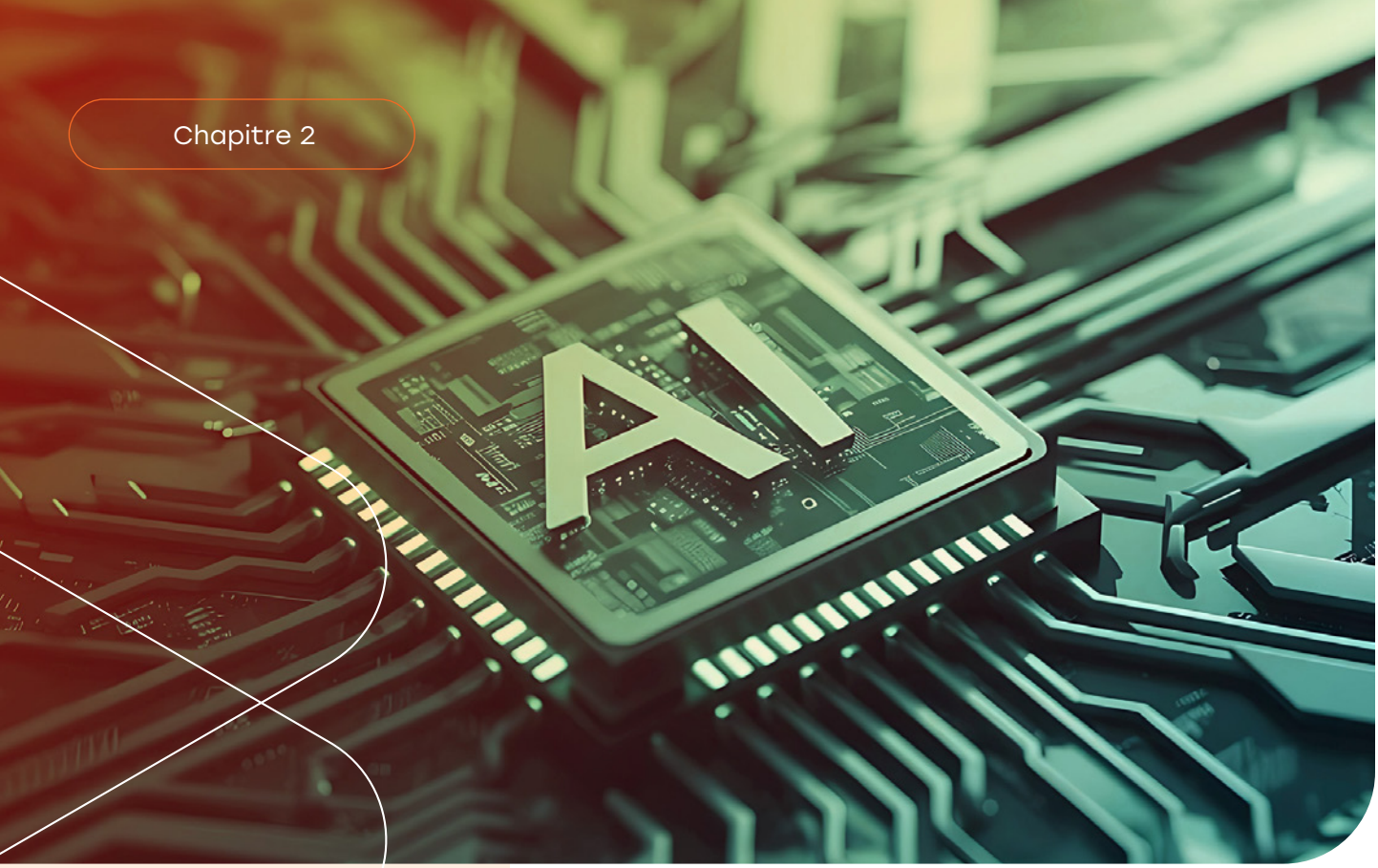
Partage excessif

Si chaque fichier d'une bibliothèque est **partagé avec tout le monde**, on peut se demander à quoi servent les révisions des accès. Le partage excessif révèle souvent des lacunes dans la structure d'une bibliothèque. Pour maintenir un niveau de risque acceptable, il est essentiel de surveiller les partages excessifs et de combler toute lacune dans les processus à mesure qu'elle se présente.

Hygiène des données

Il est difficile de maintenir un contrôle de bout en bout sur les données non structurées. Toutefois, on peut certainement commencer par les fichiers et les bibliothèques où ces fichiers sont stockés. En mettant en œuvre des contrôles appropriés dans ces bibliothèques, **conformément à leurs objectifs opérationnels** et en assurant un cycle de vie complet, nous pouvons atténuer le risque que des données inutiles ou inutilisées représentent une menace pour l'organisation en étant exposées à l'IA générative ou utilisées par celle-ci. Alors que nous continuons à générer des données chaque jour, nous devons **adopter de nouveaux comportements** qui réduisent ces risques au minimum.

Chapitre 2



Quels sont les risques ?

À l'heure actuelle, vous avez peut-être commencé à tirer des conclusions préliminaires. Quels sont les risques associés à l'utilisation d'outils puissants, comme l'intelligence artificielle dans des environnements de collaboration?

Comme de nombreux autres risques liés à la cybersécurité, ils tournent autour des données qui s'y trouvent. L'intelligence artificielle, comme toute innovation technologique, comporte des risques inédits. Cependant, comme nous l'avons vu, **l'IA générative amplifie ou accélère les risques qui existaient déjà dans les organisations bien avant le début des discussions sur l'IA.** Bien que l'efficience soit une promesse alléchante de l'IA, elle peut également entraîner des risques déstabilisants pour les entreprises mal préparées ou réticentes à l'aborder.

Quelles stratégies adopter ?

1 Commencez modestement et tirez-en parti

Ce n'est pas un **« projet »** avec une échéance précise.

À bien des égards, il peut être accablant de s'attaquer à tous ces risques, d'autant plus qu'ils existent probablement à différents niveaux depuis longtemps. Cependant, toute amélioration est un pas en avant. Sachez que ce n'est pas un « projet » avec une échéance précise. À mesure que votre organisation prend de l'expansion et que les données non structurées augmentent, **cela fait partie d'une nouvelle réalité.** Toutes les mesures que vous mettrez en œuvre seront maintenues.

Pour ce faire, vous pouvez mettre l'accent sur les partages, réduire les quantités excessives ou appliquer des contrôles appropriés aux bibliothèques où devraient se trouver vos données les plus importantes.

2 Répartissez la responsabilité

Qui génère des données et utilise ces systèmes ? → **Tout le monde**

Qui possède le contexte opérationnel des données qu'il utilise et produit ? → **Les utilisateurs finaux et les propriétaires de l'information**

Qui devrions-nous utiliser pour constater des progrès et maintenir un environnement acceptable en matière de risque à l'avenir ? → **Le plus grand nombre d'utilisateurs possible**

Fournissez à vos utilisateurs finaux du contexte et des outils simples pour les aider non seulement à tirer parti de ces outils, mais aussi à gérer les risques lorsqu'ils surviennent. Vous devrez compter sur la collaboration de chacun.

3 Simplifiez les choses

Harmonisez votre conformité et vos contrôles avec les objectifs d'affaires ou les types de données ciblés. Pour mobiliser vos utilisateurs finaux et favoriser une solide culture de cybersécurité, vous devez faire en sorte que les choses soient très simples pour tout le monde. Cela ne veut pas dire que l'arrière-plan n'est pas complexe, mais ce que vous présentez aux utilisateurs et les demandes que vous faites devraient être simples. **Vous devez établir que c'est une « habitude » pour eux.**

Vous devez faire en sorte que les choses soient **très simples pour tout le monde.**

4 C'est la nouvelle réalité

Il faut de nouveaux indicateurs pour faire le suivi. Qu'il s'agisse de la mobilisation des utilisateurs finaux, des mesures quotidiennes prises pour réduire les risques, des tendances en matière de volume de partages, des examens d'accès ou d'autres mécanismes visant à atténuer les risques, vous avez besoin de nouveaux indicateurs et de nouvelles méthodes pour en faire le suivi. Ils compléteront ou remplaceront certains des indicateurs que vous avez peut-être déjà, comme les taux d'hameçonnage, les taux de clics et l'apprentissage. Ils devraient **être communiqués aux cadres supérieurs et présentés de façon positive.**

Les points à retenir

La collaboration alimentée par l'intelligence artificielle est au cœur de nos collaborations, ce qui entraîne une augmentation soudaine des données non structurées qui nécessitent de nouvelles stratégies de sécurité. Les organisations doivent agir maintenant pour équilibrer l'innovation et la gestion des risques. Voici comment :

1

Commencer à petite échelle et stratégiquement

Mettez l'accent sur les gains rapides, comme la réduction du partage excessif de données et l'application des contrôles d'accès aux bibliothèques à risque élevé. La sécurité est un processus continu et non un projet ponctuel.

2

Habiliter les utilisateurs finaux

La sécurité n'est pas uniquement la responsabilité des TI. Il importe de fournir aux employés le contexte et les outils nécessaires pour cerner et gérer les risques liés aux données dans leurs flux de travail quotidiens. La promotion d'une bonne hygiène des données garantit que les informations sensibles sont gérées et protégées de manière appropriée.

3

Simplifier les pratiques de sécurité

Harmonisez les mesures de sécurité avec les objectifs opérationnels. Lorsque les procédures sont simples, évidentes et harmonieusement intégrées, les utilisateurs sont plus enclins à les accepter.

4

Mesurer et adapter

Surveillez les nouveaux indicateurs de sécurité, comme la participation à la révision des accès, les tendances en matière de partage des données et l'exposition aux données axées sur l'intelligence artificielle. Communiquez ces renseignements aux cadres supérieurs pour favoriser la sensibilisation et l'amélioration continue.

Que votre organisation adopte ou non pleinement l'IA, la protection des données non structurées dans Microsoft 365 est non négociable. Le moment est venu d'agir – chaque mesure que vous prenez renforce votre cyberrésilience, améliore **l'hygiène des données** et protège votre entreprise contre les menaces émergentes.

À propos de WeActis

WeActis révolutionne la cybersécurité en transformant les employés en contributeurs actifs, et en faisant de la sécurité une responsabilité collective, et non pas un fardeau informatique.

Intégrée à Microsoft Teams, l'application **WeActis** encourage les employés à prendre des actions sécuritaires simples permettant de diminuer les risques de leur organisation en moins de cinq minutes par semaine.

En favorisant une meilleure hygiène des données dans Microsoft 365, en renforçant la gouvernance et en simplifiant la gestion des risques, **WeActis** permet aux organisations de bâtir une culture de cyberrésilience. Cela se traduit par une meilleure conformité, une réduction de l'exposition des données et des améliorations en sécurité, à la fois mesurables et durables.

info@weactis.com
1-833-666-3282
WeActis.com

WeActis
par Mondata