

L'observabilité des données démystifiée : des fondamentaux aux usages avancés

Faites de l'observabilité un pilier stratégique de votre entreprise. Détectez les problèmes en amont et corrigez-les avant que des données défectueuses n'entraînent de mauvaises décisions.

Table des matières

- 2 Synthèse exécutive : la santé des données n'est plus négociable
- 3 Comprendre les bases : qu'est-ce que l'observabilité des données ?
- 3 L'enjeu business : garantir la confiance dans les données
- 4 Le facteur humain : qui utilise l'observabilité des données et pourquoi
- 5 Le rôle technique de l'observabilité des données
- 6 4 cas d'usage techniques : soutenir la dimension technologique de l'observabilité
- 8 7 agents d'observabilité des données et leurs missions
- 8 MCP et connectivité agentique : intégrer l'observabilité aux assistants IA
- 9 6 étapes pour moderniser votre approche de l'observabilité des données
- 9 À propos d'Actian

Synthèse exécutive :

la santé des données n'est plus négociable

Les dirigeants d'entreprise sont devenus experts dans la mesure de la performance. Les tableaux de bord sont omniprésents et les indicateurs couvrent tous les domaines. Le problème, c'est que beaucoup de ces tableaux de bord ressemblent à un compteur de vitesse sans jauge de carburant. Vous voyez à quelle vitesse vous avancez... jusqu'au moment où tout s'arrête.

L'observabilité des données, c'est justement cette jauge et ce diagnostic moteur dont les entreprises ont besoin, mais qu'elles n'ont pas toujours. Elle garantit que les données qui alimentent l'IA, l'automatisation et la prise de décision sont fiables et à jour. Sans contrôles de fraîcheur, par exemple, des prévisions peuvent être retardées ou basées sur des chiffres obsolètes, créant des risques évitables dans les décisions quotidiennes.

Avec l'IA agentique qui entre en scène, l'exigence en matière de qualité des données s'élargit encore. Il ne suffit plus de détecter les problèmes après coup. Il faut des systèmes capables de valider les données dès leur arrivée, d'expliquer leurs transformations dans un langage métier clair et de corriger les anomalies avant que des données peu fiables ne produisent des résultats tout aussi peu fiables.

C'est dans cette dynamique que l'observabilité évolue vers des agents de fiabilité des données. Ces agents accélèrent l'identification des causes racines et assurent un flux continu de données fiables vers les analystes, les domaines métiers et les systèmes d'IA qui en dépendent.

Comprendre les bases : qu'est-ce que l'observabilité des données ?

L'observabilité des données est la capacité à comprendre l'état et la santé des données au sein de vos systèmes en mesurant des signaux clés comme la qualité, la fiabilité et le lignage des données. Elle permet de répondre rapidement à trois questions essentielles :

- Les données sont-elles erronées ?
- Qu'est-ce qui a dysfonctionné, et où ?
- Comment corriger le problème ?

Une véritable approche d'observabilité dépasse le simple tableau de bord ou l'outil ponctuel. Elle propose une vision globale et automatisée pour surveiller les données, détecter les anomalies en amont et réduire le délai entre l'incident et sa résolution, à travers les pipelines, les tables et les processus métiers critiques.

L'exigence en matière de qualité des données ne cesse d'augmenter. Les organisations investissent massivement dans l'IA et l'analytique, mais les projets échouent lorsque la qualité des données et les mécanismes de contrôle ne sont pas à la hauteur ou ne sont pas jugés fiables. Sans vision holistique de la santé des données, le récit reste incomplet et les insights limités.

Les données doivent être fiables pour que les résultats métiers, y compris les cas d'usage IA, le soient également.



Figure 1.

Un pipeline supprime silencieusement des colonnes d'une table. Le tableau de bord dérive sans information complète. Des décisions sont prises sur la base de ce tableau de bord devenu peu fiable. Les prévisions sont manquées. La direction perd confiance dans les données.

L'enjeu business : garantir la confiance dans les données

L'observabilité des données ne consiste pas à éteindre des incendies. Elle vise à soutenir des décisions plus rapides, de meilleures expériences client et une gouvernance renforcée, parce que les équipes peuvent trouver et activer des données auxquelles elles font confiance.

Des données de qualité sont indispensables pour l'IA et de nombreux autres cas d'usage. Gartner estime que d'ici 2025, au moins 30 % des projets d'IA auront été abandonnés après la preuve de concept en raison d'une mauvaise qualité des données, de contrôles de risque insuffisants, de coûts croissants ou d'une valeur métier peu claire. Gartner souligne également que 93 % des équipes ne détectent pas les problèmes de données avant qu'ils n'impactent l'activité.

L'observabilité des données soutient des cas d'usage concrets tels que :

- **Continuité opérationnelle.** Réduire les interruptions lorsque des systèmes en amont évoluent.
- **Protection de l'expérience client.** Limiter les recommandations erronées, les erreurs de tarification ou les problèmes d'expédition.
- **Gestion des risques et conformité.** Bénéficier d'une meilleure visibilité sur le traitement des données sensibles et leur lignage.
- **Confiance dans l'IA.** Réduire la surcharge liée aux workflows automatisés et fiabiliser les résultats produits.

Le coût des problèmes de qualité augmente à mesure que les données circulent vers l'aval. Par exemple, une simple modification de schéma en amont peut provoquer un effet domino rapide (Figure 1).

L'observabilité des données interrompt cette chaîne d'événements dès son origine, en mettant en évidence le problème là où il apparaît, avant qu'il n'affecte les usages en aval.

Le facteur humain : qui utilise l'observabilité des données et pourquoi

L'observabilité des données ne fonctionne que si elle répond à des problèmes concrets. Les profils qui l'utilisent et les enjeux auxquels elle répond incluent :

Data engineers Responsables de la construction et de la maintenance des infrastructures et des pipelines de données qui alimentent les systèmes d'observabilité. Ils ont besoin d'une supervision complète des pipelines afin d'identifier et de corriger de manière proactive les problèmes avant qu'ils n'affectent les utilisateurs en aval et les processus métiers.

Data analysts Spécialistes du traitement et de l'analyse des données, ils dépendent de données observables et de haute qualité pour produire des rapports et des analyses fiables. Ils doivent avoir la certitude que les données sont exactes, fiables et à jour, sans devoir remettre en question leur qualité en permanence.

Data scientists Ils utilisent des techniques d'analytique avancée et s'appuient sur des données gouvernées et de qualité pour entraîner et déployer leurs modèles. Ils ont besoin de visibilité sur la qualité des données et leur lignage afin de faire confiance aux jeux de données alimentant leurs modèles et de comprendre les baisses de performance liées à des problèmes de données.

Ingénieurs DevOps Ils gèrent l'infrastructure technique qui supporte les plateformes d'observabilité et assurent la fiabilité des systèmes. Ils ont besoin de mécanismes de surveillance et d'alerting pour maintenir l'infrastructure des opérations data avec un minimum d'interruptions et des performances optimales.

Développeurs Business Intelligence (BI) Ils créent des tableaux de bord et des rapports reposant sur des sources de données fiables et observables. Ils ont besoin d'outils d'observabilité intégrés à leurs plateformes de reporting pour garantir l'exactitude des indicateurs et alerter les utilisateurs métiers lorsque des problèmes de qualité peuvent affecter leurs rapports.

Responsables conformité Ils veillent au respect des exigences réglementaires en matière de gestion des données. Ils ont besoin d'un lignage complet et de pistes d'audit détaillées, jusqu'au niveau des champs, afin de démontrer la conformité et de répondre rapidement aux demandes liées à la gouvernance des données.

Responsables IT Ils supervisent l'infrastructure technologique et doivent avoir une visibilité sur la performance et les coûts des solutions d'observabilité. Ils recherchent des plateformes capables de fournir des indicateurs clairs de retour sur investissement et d'optimiser les coûts d'infrastructure tout en maintenant un haut niveau de service.

Chief Data Officers (CDO) Ils pilotent la stratégie data à l'échelle de l'entreprise. Ils ont besoin de métriques globales sur la qualité, l'usage et l'impact métier des données pour démontrer la valeur des investissements et identifier les axes d'amélioration stratégique.

La confiance dans les données ne doit pas être centralisée dans un seul département. Elle doit être distribuée afin que les utilisateurs métiers puissent poser et résoudre des questions liées à la fiabilité sans attendre qu'une équipe spécialisée intervienne. Les interfaces en langage naturel et les agents d'observabilité spécialisés contribuent à réduire les goulots d'étranglement et à accélérer le temps de résolution.

Le rôle technique de l'observabilité des données

L'observabilité contribue à garantir que les données sont exactes et adaptées à leur usage dans l'ensemble des systèmes et des processus. Elle consiste à surveiller, gérer et maintenir les données de manière à assurer leur qualité, leur disponibilité et leur fiabilité.

Atteindre les objectifs métiers de l'observabilité repose sur quatre capacités techniques clés :

1. **Supervision des données.** Voir ce qui se passe en temps réel. C'est essentiel pour les applications nécessitant une action immédiate, comme les systèmes de détection de fraude ou les stratégies de tarification dynamique.
2. **Détection et résolution automatisées.** Identifier les anomalies et réagir rapidement. Les solutions modernes d'observabilité doivent gérer de grands volumes de données et détecter efficacement des anomalies complexes. Cela est particulièrement critique dans des secteurs comme le retail, la finance ou les télécommunications, où la détection et la correction automatiques sont nécessaires pour maintenir l'efficacité opérationnelle et la satisfaction client.
3. **Qualité et accessibilité des données.** Garantir que les données sont exploitables et fiables. La qualité des insights est directement corrélée à la qualité des données collectées et stockées dans vos systèmes. Des données exactes, propres et facilement accessibles sont indispensables pour l'analyse, le reporting et l'entraînement de modèles d'IA.
4. **Capacités avancées.** Exploiter la détection d'anomalies pilotée par l'IA et l'analyse automatisée des causes racines. Les plateformes d'observabilité évoluent pour répondre aux nouveaux défis en s'appuyant sur le machine learning afin d'identifier des schémas inhabituels.

Une solution d'observabilité des données offre une visibilité complète sur les flux et dépendances, tout en facilitant le diagnostic rapide et la résolution des incidents. L'ensemble de ces opérations est réalisé directement au sein de l'environnement d'observabilité.




4 cas d'usage métier : où l'observabilité s'impose dans les secteurs d'activité

L'observabilité des données soutient directement des objectifs métiers critiques dans tous les secteurs. Les organisations l'utilisent pour générer des résultats mesurables.

1. Services financiers : identifier les risques et la fraude

Renforcez l'évaluation des risques et la prévention de la fraude en détectant les anomalies dans les schémas de transactions. Validez également la fraîcheur des flux de marché et soutenez la conformité réglementaire.




Exemples

-  **Banque.** Détecter les anomalies dans les flux de transactions afin d'identifier des fraudes potentielles, tout en assurant la conformité aux réglementations anti-blanchiment grâce à un suivi complet du lignage des données.
-  **Assurance.** Surveiller les données de sinistres et les délais de traitement pour repérer des comportements frauduleux et garantir la fiabilité des modèles d'évaluation des risques.
-  **Gestion d'actifs.** Contrôler la fraîcheur et la qualité des données afin que les décisions de gestion de portefeuille reposent sur des informations exactes et actualisées.

2. Santé : sécurité des patients et excellence opérationnelle

Lorsque les données issues des dossiers médicaux électroniques (EHR) ou des admissions sont incomplètes, retardées ou incohérentes, les conséquences peuvent être graves. L'observabilité permet une surveillance proactive de la qualité des données patients, des données issues des dispositifs médicaux et de l'efficacité des workflows cliniques afin d'améliorer les résultats médicaux et la performance opérationnelle.

Exemples

-  **Hôpitaux.** Surveiller la qualité des données EHR pour garantir l'exactitude des informations patients, des dosages médicamenteux et des recommandations de traitement.
-  **Industrie pharmaceutique.** Suivre l'intégrité des données d'essais cliniques et la conformité réglementaire tout au long du développement des médicaments.
-  **Dispositifs médicaux.** S'assurer que les données IoT issues des équipements de surveillance des patients restent exactes et disponibles pour soutenir des décisions critiques.

3. Retail et e-commerce : expérience client

Les catalogues produits, la disponibilité des stocks, les recommandations et la tarification dépendent tous de pipelines fiables. L'observabilité permet d'éviter les erreurs silencieuses qui génèrent de la frustration client. C'est l'équivalent métier d'un contrôleur aérien pour votre système d'information : suivre en continu les flux, détecter les anomalies en amont et intervenir avant que de petits incidents ne deviennent des défaillances visibles.




Exemples

-  **Plateformes e-commerce.** Surveiller la qualité des données catalogue, la précision des stocks et la performance des algorithmes de recommandation afin d'améliorer la satisfaction client et d'éviter des pertes de ventes.
-  **Retail omnicanal.** Garantir la cohérence des données clients entre les canaux en ligne, mobile et en magasin pour offrir une expérience fluide.
-  **Chaîne d'approvisionnement.** Suivre la qualité des données d'inventaire et les indicateurs de performance fournisseurs afin d'optimiser les niveaux de stock et d'éviter les ruptures ou les surstocks.

4. Industrie manufacturière : optimisation de la production et contrôle qualité

Obtenez des insights sur la qualité des données de production, la performance des équipements et l'efficacité de la chaîne d'approvisionnement afin d'optimiser les processus industriels et garantir que les produits respectent les spécifications attendues.

Exemples

-  **Automobile.** Surveiller les données issues des capteurs des lignes de production pour assurer le respect des standards qualité et anticiper les besoins de maintenance.
-  **Aéronautique.** Suivre la traçabilité des composants et les indicateurs qualité tout au long du processus de fabrication afin de garantir la sécurité et la conformité réglementaire.
-  **Biens de consommation.** Analyser les tendances des données de production pour optimiser les lots et réduire les déchets tout en maintenant des standards élevés de qualité.

Surveiller en continu les pipelines pour contrôler la fraîcheur, l'exactitude et les changements inattendus permet de faire confiance aux données, d'agir plus vite et d'éviter les coûts cachés liés à une mauvaise qualité avant qu'ils n'impactent la performance financière.

4 cas d'usage techniques : soutenir la dimension technologique de l'observabilité

Vous êtes-vous déjà demandé comment les stratégies métier deviennent une réalité numérique ? Ces cas d'usage lèvent le voile sur la puissance technique qui se cache derrière l'observabilité des données. Ils montrent comment l'observabilité transforme l'intention business en données fiables, en insights exploitables et en avantage concurrentiel.

1. Supervision des pipelines de données et reprise automatisée

Surveillez en continu des pipelines ETL/ELT complexes et automatisez la reprise après certains types d'incidents. Cela garantit la résilience nécessaire aux opérations data critiques.

Capacités clés

- **Supervision de la santé des pipelines** avec alertes automatiques en cas d'échec de traitement, de dégradation des performances ou de problèmes de qualité des données.
- **Mécanismes de reprise automatisés** pour les incidents courants, incluant des logiques de relance et des sources de données de secours.
- **Recommandations d'optimisation des performances** basées sur les historiques d'exécution et l'utilisation des ressources.

2. Scoring de la qualité des données et gestion des SLA

Mettez en place un scoring automatisé de la qualité des données afin d'évaluer leur état par rapport aux SLA définis et aux exigences métiers. Ainsi, le niveau de qualité "acceptable" devient mesurable, contrôlable et visible.

Capacités clés

- **Scoring automatisé de la qualité** avec pondérations configurables selon différentes dimensions.
- **Suivi et reporting des SLA** via des tableaux de bord affichant la performance par rapport aux indicateurs établis.
- **Alertes proactives** lorsque les scores de qualité passent sous les seuils acceptables.
- **Analyse des tendances historiques** pour identifier les schémas de dégradation.

3. Gestion des incidents et analyse des causes racines

Détectez les incidents, accélérez le triage et facilitez l'analyse des causes racines. Réduisez le délai moyen de résolution (MTTR) lorsque des problèmes de données surviennent.

Capacités clés

- **Corrélation intelligente des incidents** regroupant les problèmes liés et mettant en évidence les causes probables.
- **Analyse automatisée de l'impact** montrant quels systèmes en aval et quels utilisateurs sont affectés.
- **Gestion collaborative des incidents** avec outils de communication intégrés et automatisation des runbooks.
- **Analyse post-incident** avec génération automatique de rapports et recommandations pour éviter les récidives.

4. Lignage des données et analyse d'impact

Offrez un lignage de bout en bout et une analyse d'impact complète afin de comprendre les dépendances, gérer les changements en toute sécurité et relier les actifs techniques à leur signification métier.

Capacités clés

- **Découverte automatisée du lignage** dans des écosystèmes de données complexes et multi-éditeurs.
- **Suivi du lignage** au niveau colonne pour visualiser en détail les transformations et dépendances jusqu'au champ.
- **Analyse d'impact des changements**, permettant d'anticiper les effets en aval d'une modification en amont.
- **Intégration d'un glossaire métier** reliant le lignage technique aux termes et définitions business.

Les organisations s'appuyaient autrefois sur des solutions d'observabilité pour être alertées lorsqu'un pipeline était défaillant. La nouvelle attente est différente : détecter les problèmes de qualité en amont, les expliquer clairement et aider à les corriger rapidement.

Passez à l'étape suivante : l'observabilité pour les workflows modernes

Les agents d'observabilité des données adoptent une approche moderne de la qualité en validant les données à la source plutôt qu'en aval. Les stratégies traditionnelles de qualité détectent souvent les problèmes une fois que les données ont déjà traversé les pipelines et atteint les utilisateurs.

En déplaçant le contrôle en amont, dès la mise à disposition des données, vous évitez que des données de mauvaise qualité n'alimentent des tableaux de bord, des applications ou des workflows IA.

7 agents d'observabilité des données et leurs missions

À mesure que les attentes évoluent, passant de la simple visibilité des incidents à leur prévention, l'observabilité doit gagner en capacités. C'est là qu'interviennent ses agents d'observabilité des données.

Conçus pour fonctionner ensemble tout au long du cycle de vie des données, de la détection à la résolution, ces agents assurent la livraison et le maintien de données fiables. Actian propose sept agents spécialisés, chacun doté de compétences ciblées :

1. **Agent de validation.** Recommande et applique des règles de validation telles que les contrôles de schéma, de complétude, de cohérence et de fraîcheur dès l'ingestion. Il empêche les "mauvaises données" d'entrer dans les flux en aval.
2. **Agent de diagnostic des incidents.** Détecte les anomalies et explique les causes racines probables en langage clair, comme une modification de flux en amont, un échec de traitement ou une hausse soudaine des valeurs nulles.
3. **Agent de lignage.** Cartographie les dépendances pour visualiser le périmètre d'impact d'un actif de données défaillant sur les tables, tableaux de bord et modèles concernés.

4. **Agent d'insight data.** Traduit les signaux de qualité en recommandations compréhensibles par les parties prenantes, en expliquant ce qui a changé et les impacts associés. Il renforce la confiance grâce à des insights clairs et actionnables.
5. **Agent d'orchestration.** Automatise les étapes de résolution et s'intègre aux workflows de ticketing et d'alerting afin de boucler le cycle entre détection et remédiation.
6. **Agent de routage.** Alertes les bons responsables selon le domaine, la sévérité et l'impact. Il évite que les incidents restent bloqués dans des canaux sans propriétaire clairement identifié.
7. **Agent d'assistance.** Guide les utilisateurs dans la configuration et le dépannage via une interface en langage naturel. Il facilite la création de moniteurs, la compréhension des incidents et l'amélioration continue de la qualité.

Ensemble, ces agents font évoluer l'observabilité d'un modèle de surveillance réactive vers des actions proactives et autonomes, au service d'une fiabilité durable et maîtrisée.

MCP et connectivité agentique : intégrer l'observabilité aux assistants IA

Les équipes data travailleront de plus en plus avec des assistants et agents IA. Un serveur conforme au Model Context Protocol (MCP) permet aux agents IA d'accéder au contexte d'observabilité sans intégrations spécifiques.

Une couche MCP permet aux agents de récupérer des données validées ainsi que les métadonnées de qualité nécessaires à une prise de décision éclairée. On peut voir le MCP comme le "USB-C" de la connectivité entre l'IA et les outils : une interface standardisée qui réduit les connexions spécifiques et accélère une adoption sûre des assistants IA et des frameworks d'agents.

Pourquoi Actian Data Observability se distingue

La plupart des outils d'observabilité sont optimisés pour une plateforme unique et peinent lorsque les données sont réparties entre des composants lakehouse, des formats de tables ouverts et plusieurs environnements cloud. Actian adopte une approche différente, en proposant une observabilité conçue pour la diversité, l'ouverture et la fiabilité in situ à travers l'écosystème lakehouse.

Les différenciateurs Actian incluent :

- **Couverture Microsoft Fabric / OneLake. Surveillez les données sur place.** Profilage et monitoring directement au niveau OneLake afin que les utilisateurs Microsoft Fabric bénéficient d'une fiabilité continue sans déplacer les données.
- **Observabilité indépendante du catalogue pour les architectures lakehouse ouvertes.** Actian offre une couverture étendue quel que soit le catalogue standardisé.
- **Conçue pour des architectures fédérées et non pour un modèle mono-plateforme.** Actian prend en charge des architectures ouvertes et fédérées où les données résident sur plusieurs plateformes et clouds, afin que la fiabilité ne dépende pas d'un environnement unique.

Si votre lakehouse est multi-plateforme, multi-format et en constante évolution, Actian garantit une observabilité cohérente. La confiance ne dépend pas de l'endroit où les données sont stockées.

5. **Renforcez la confiance des analystes.** Décentralisez la confiance afin que les utilisateurs métiers puissent interpréter l'état de santé des données en toute autonomie.
6. **Rendez les agents IA sensibles à la qualité via MCP.** Permettez aux workflows agentiques de vérifier le contexte de qualité avant d'agir.

L'observabilité des données a commencé comme un outil de visibilité. Elle évolue désormais vers une approche plus opérationnelle : la fiabilité des données, capable de soutenir l'automatisation et la prise de décision agentique sans compromettre la confiance.

Les agents d'observabilité des données accompagnent cette évolution en combinant validation proactive à l'ingestion, détection et remédiation autonomes, explications en langage clair et workflows connectés via MCP. Ainsi, humains et agents IA opèrent dans un contexte partagé et fiable.

Si votre organisation prend l'IA agentique au sérieux, la question n'est pas de savoir si vous avez besoin d'observabilité. Elle est de savoir si votre observabilité est prête à soutenir l'autonomie avec des garde-fous adaptés.

À propos d'Actian

Actian aide les entreprises à gérer et gouverner leurs données en toute confiance, à grande échelle. Les organisations font confiance aux solutions de gestion et d'intelligence des données d'Actian pour maîtriser des environnements data complexes et accélérer la mise à disposition de données prêtes pour l'IA. Conçues pour être flexibles et fiables, ces solutions s'intègrent facilement à tous types d'environnements : sur site, dans le cloud ou en mode hybride. Actian est la division Data de HCLSoftware. Pour en savoir plus : www.actian.com.

6 étapes pour moderniser votre approche de l'observabilité des données

Si vous modernisez votre observabilité des données en vue d'un avenir orienté IA, privilégiez le progrès plutôt que la perfection. Ces étapes peuvent vous guider :

1. **Commencez par les actifs critiques.** Priorisez les données qui soutiennent le chiffre d'affaires, la gestion des risques, la sécurité des patients ou l'expérience client.
2. **Définissez des signaux de confiance alignés sur les résultats.** Assurez la fraîcheur, la complétude, la stabilité des schémas et la stabilité des distributions en les reliant à des SLA concrets.
3. **Automatisez la détection et l'explication dès le départ.** Réduisez le temps nécessaire pour comprendre les incidents et utilisez-le comme levier de retour sur investissement.
4. **Bouclez la chaîne avec l'orchestration.** La détection sans action n'améliore pas la qualité. Intégrez le routage, le ticketing et les workflows de remédiation.