

LIVRE BLANC

COMMENT BÉNÉFICIER DE DONNÉES DE TEST SAP DE QUALITÉ ET À JOUR ?



*Livre blanc écrit par Mathilde Régnier-Dulout,
Directrice Générale adjointe chez Invarture*

De nos jours, les données dont dispose une organisation sont souvent comparées à du pétrole brut : une vraie richesse dont l'utilisation directe est limitée ; il faut passer par des étapes successives avant de bénéficier de son potentiel. Comme le pétrole doit subir un raffinage avant de donner carburants, matières plastiques, engrais, détergents, ..., de même les données doivent subir une série de traitements avant d'être réellement utilisables : vérification, consolidation, anonymisation, validation, ...

Si la mise en place d'ERP tels que SAP a grandement amélioré la qualité des données et donc leur disponibilité en production, il n'en reste pas moins que la qualité de ces données dans les environnements hors production peut demeurer beaucoup plus problématique.

**« L'IMPACT FINANCIER MOYEN DE DONNÉES DE QUALITÉ MÉDIOCRE SUR LES ORGANISATIONS EST DE 9,7 M\$/AN »
- GARTNER, 2017¹**

Si l'importance stratégique de la qualité des données est maintenant reconnue, dans la majorité des cas seule la qualité des données en production est au centre des préoccupations. Les avantages amenés par des données de qualité dans l'atteinte des objectifs Métier (et les risques résultant de la médiocrité des données ...) sont bien connus des dirigeants et des efforts sont souvent consentis pour l'amélioration de cette qualité.

Il n'en est toutefois pas toujours de même pour les données de test. En effet, si dans les environnements de test, la qualité (et la quantité !) des données disponibles devrait constituer une des préoccupations majeures de toute organisation, l'expérience montre que ce n'est généralement pas le cas.

« EN MOYENNE, 15 % DE L'ENSEMBLE DES DÉFAUTS LOGICIELS SONT IMPUTABLES AUX DONNÉES » - DELPHIX, STATE OF TEST DATA MANAGEMENT, 2017²

¹<https://www.gartner.com/smarterwithgartner/how-to-create-a-business-case-for-data-quality-improvement/>

²<https://www.delphix.com/sites/default/files/2017-07/State-Of-Test-Data-Management-Report.pdf>

QU'EST-CE QUE LA QUALITÉ D'UNE DONNÉE ?

Pour estimer la qualité des données de test, les indicateurs clés utilisés sont l'intégrité, l'âge, la pertinence par rapport au secteur et la sécurité :

- **Intégrité**

Des données de test médiocres manqueront souvent de l'intégrité propre aux données de production, elles seront dépourvues des relations exactes et cohérentes existant dans les systèmes (par ex. ECC : Produit Bon de Commande Vendeur) et ne seront pas reliées de manière cohérente entre systèmes (par ex. ECC CRM, ECC - BW).

- **Âge des données**

Il revêt deux aspects :

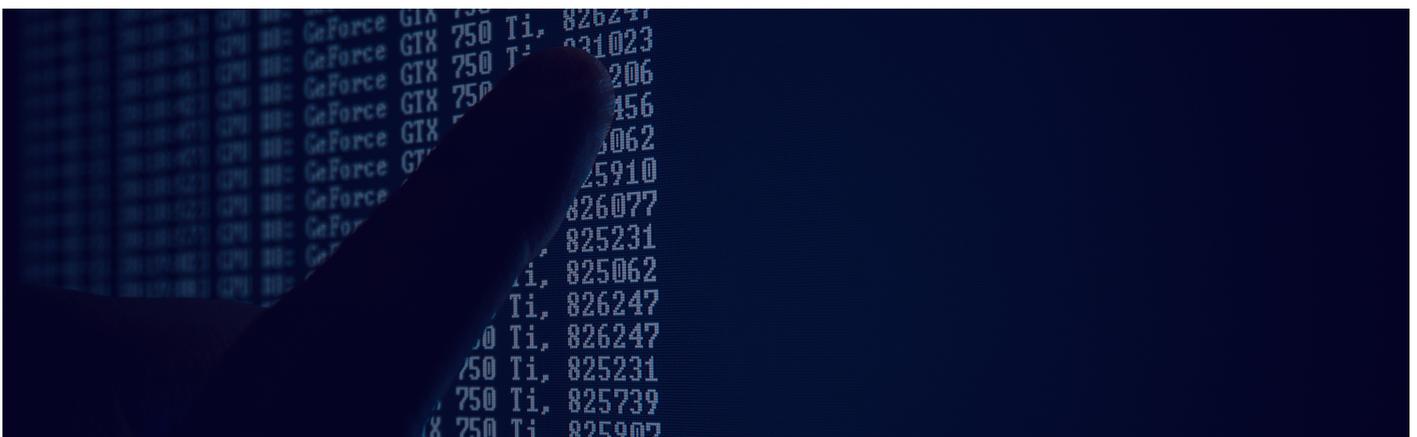
- La plupart du temps, les 10 % des données les plus récentes des systèmes sont importants pour les tests; les 90 % restant utilisent de l'espace dans l'environnement de test et ont peu d'utilité ; des périodes de 3 à 6 mois satisfont la majeure partie des besoins.
- Le deuxième point concerne la fraîcheur des données de test par rapport à celles de la production. Les clients SAP recourant à un rafraîchissement annuel de leur environnement QA utilisent pendant 9 mois de l'année des données de test vieilles de plus de 3 mois.

- **Pertinence**

Le cycle de rafraîchissement est-il aligné avec les besoins du secteur ? Par exemple, les systèmes SAP Retail ou Utilities manipulent de gros volumes de données et ont des cycles transactionnels très courts, créant ainsi un besoin en données récentes et nombreuses dans les systèmes de test.

- **Sécurité**

Les données de test devraient être semblables à celles de production, mais n'ont pas besoin d'être identiques (ce qui serait un casse-tête en matière de sécurité et de gouvernance). Les informations de test de nature personnelle et sensible doivent être anonymisées afin de garantir que les données et la réputation de l'organisation ne sont pas mises en danger.



Mais quand la qualité des données de test est-elle primordiale ? Tout d'abord en support quotidien de la production, pour l'analyse et la résolution de problèmes. Mais également pour tout ce qui est montées de version et patching (EHP8, S/4, etc.), ou pour l'activation de nouvelles fonctionnalités ou configurations Métier afin de répondre à un changement de processus.

Enfin, pour la maintenance de masse des Master Data, que ce soit par exemple dans le cadre de modification des conditions tarifaires, ou des opérations nécessaires lors de la consolidation d'un ERP, ou de la scission d'un système ERP suite à une cession, séparation ou restructuration Métier.

« LA PART DES DÉPENSES IT TOTALES CONSACRÉE À L'ASSURANCE QUALITÉ ET AUX TESTS EST DE 26 %, CE QUI CORRESPOND À UNE PROPORTION RAISONNABLE DE 25 % QUI CARACTÉRISE UN BON ÉQUILIBRE ENTRE LES ACTIVITÉS DE CRÉATION DE FONCTIONNALITÉS ET CELLES DE VALIDATION. »
CAPGEMINI, WORLD QUALITY REPORT 2017-18³

Pour alimenter en données leurs environnements de test, beaucoup d'organisations se contentent de réaliser de temps en temps des rafraîchissements de masse, en copiant le plus souvent la totalité des données de production, ou encore elles « bricolent » des jeux de test à la main, avec tous les risques d'erreurs de saisie et la faible volumétrie que cela implique.

Dans un cas, la qualité des données est assurée, mais la rapidité de mise à disposition est problématique ; dans l'autre, les données de test peuvent être disponibles beaucoup plus rapidement (quoique ...), mais la volumétrie et la variété des données ne sont pas suffisantes pour refléter ce qui se passe en production et donc pour assurer la qualité des tests.



« IL N'EST PAS RARE POUR DES ORGANISATIONS DE MAINTENIR DES ENVIRONNEMENTS HORS PRODUCTION AVEC PLUS DE 90 % DE DONNÉES REDONDANTES. » DELPHIX, STATE OF TEST DATA MANAGEMENT, 2017⁴

³ <https://www.cappgemini.com/service/world-quality-report-2017-18/>

⁴ <https://www.delphix.com/sites/default/files/2017-07/State-Of-Test-Data-Management-Report.pdf>



Il semble donc qu'il faille choisir entre qualité et rapidité de mise à disposition ... Quels sont alors les impacts de chacun de ces choix ? Lorsque les équipes procèdent à une copie de l'ensemble des données de production dans les environnements de test, les volumes manipulés sont tels qu'il faut souvent plusieurs jours voire semaines avant de disposer des données. Et de pouvoir reprendre les tests. Les projets sont pendant ce temps mis en attente, et le taux d'utilisation des équipes de test réduit.

Par ailleurs, le choix d'une telle solution entraîne des coûts réseau ou disque hors production élevés, ainsi que des coûts de maintenance des dispositifs hors production égaux à ceux de la production. Par ailleurs, une dépendance inter-équipes complexe apparaît à cette occasion (l'équipe Basis a besoin de l'équipe Infra pour l'allocation de disque, l'équipe Fonctionnelle a besoin de l'équipe Basis pour la copie complète du système, aucune équipe n'est prééminente). L'accès aux environnements hors production des équipes fonctionnelles et de test devra également être restreint en cas d'absence d'encodage des données, afin de garantir la sécurité et la confidentialité nécessaires des données.

Devant l'importance de l'impact de ces copies complètes de données, certaines organisations préfèrent ne procéder que rarement à des rafraîchissements. Les équipes de test se retrouvent alors à manipuler des données qui ne reflètent plus la réalité Métier opérationnelle, et qui donc compromettent la valeur et la pertinence des modifications apportées aux systèmes (cf. page 2 « Qu'est que la qualité d'une donnée ? »).

Une autre solution est de « bricoler » manuellement des sous-ensembles de données afin de réduire la quantité de données ; l'Âge ainsi que la Sécurité des données sont alors (plus ou moins) respectés, au détriment toutefois de leur Intégrité et de leur Pertinence. Par ailleurs, la validité des tests peut dans ce cas être contestable, la qualité de ces tests ne pouvant être au mieux qu'équivalente à celle des données générées manuellement...

<< SEULES 1/4 DES ORGANISATIONS RECOURENT À DES OUTILS DE MASQUAGE DES DONNÉES DE TEST >> - DELPHIX, STATE OF TEST DATA MANAGEMENT, 2017⁵

⁵ <https://www.delphix.com/sites/default/files/2017-07/State-Of-Test-Data-Management-Report.pdf>

La suite **DSM (Data Sync Manager)** proposée par Invarture répond de façon rapide et économique à ces problématiques.

L'outil **Client Sync** de la suite permet de réaliser des copies partielles de mandant, en procédant à un découpage temporel et/ou organisationnel des données de production pour les copier dans les environnements hors production.

L'outil **Object Sync** permet de sélectionner des données au niveau d'un objet de gestion ou d'un flux transactionnel (allant d'un objet individuel jusqu'à un ensemble d'objets complet), puis de les copier dans un système hors production ; les données sont transférées de manière cohérente, tous les liens aux données associées sont conservés et l'intégrité des données est entièrement garantie. Par ailleurs, Object Sync offre une protection des données sensibles grâce à des fonctions de chiffrement ou de conversion utilisées principalement par les utilisateurs fonctionnels.

L'outil **Data Secure** permet également de chiffrer les données sensibles ; il est toutefois plutôt utilisé par l'équipe d'administration du système et sert au chiffrement en masse de mandants entiers. Toutefois, les routines de base utilisées pour la génération de nouvelles valeurs sont identiques dans Object Sync et Data Secure. La convivialité et l'aspect des données rendues anonymes restent donc cohérents.

L'outil **System Builder**, quant à lui, offre la possibilité de créer un nouveau système « coquille » hors production prêt à l'emploi, avec un Repository correspondant exactement à celui du système de production, mais sans mandant productif. Les objets du Repository et la configuration inter-mandants du nouveau système sont synchronisés avec le système de production afin de bénéficier en permanence de données à jour dans un système mobilisant beaucoup moins d'espace disque.



LA TECHNOLOGIE EST FRANCHEMENT CE QU'IL Y A DE PLUS SIMPLE À ACQUÉRIR. IL EST VRAIMENT, VRAIMENT IMPORTANT QUE LES ORGANISATIONS METTENT EN PLACE AU PRÉALABLE DES BENCHMARKS ET UNE COMPRÉHENSION DE LEUR POLITIQUE ET LES CONFRONTENT À LA TECHNOLOGIE AFIN DE S'ASSURER QUE CELLE-CI POURRA ACCOMPAGNER CE QUI EST NÉCESSAIRE >>
MICHELE GOETZ, ANALYSTE PRINCIPAL CHEZ FORRESTER⁶

⁶ https://www.forbes.com/forbesinsights/pitney_bowes_data_quality/index.html

Le besoin en données de test récentes et qualitatives devient par ailleurs de plus en plus essentiel dans les organisations, avec des temps de mises en production logicielles devant être sans cesse raccourcis et le recours aux méthodes Agile ou DevOps. Selon l'étude Delphix, la fourniture aux équipes de test des « bonnes » données implique la prise en compte des points suivants :

- Mise à disposition des données : réduction du temps de mise à disposition des données de test.
- Qualité des données : respect des exigences de qualité de données (cf. encadré).
- Sécurité des données : minimisation des risques tout en conciliant les besoins en rapidité.
- Coûts d'infrastructure : réduction des coûts de stockage et d'archivage des données de test.

Grâce à la suite DSM, le volume de données de test peut être réduit afin de cibler les besoins, d'assurer l'intégrité à la fois dans et à travers les différents systèmes ; les données de test peuvent également être facilement anonymisées afin de ne pas divulguer des données sensibles en dehors du système de production. La qualité et la fraîcheur des données de test sont ainsi assurées, tout en minimisant les coûts d'infrastructure.

Qui sommes-nous ?

Invature est une société française proposant des solutions logicielles qui aidant les ETI et grands comptes européens à optimiser leur système d'information et notamment leur ERP (SAP).

Notre expertise s'articule autour de 3 domaines spécifiques :

- **DevOps et livraison continue**
- **sécurité et gestion des données**
- **transformation digitale et expérience utilisateurs (UX)**

Invature
Beyond your ERP



www.invature.com



info@invature.com