



Livre blanc 

---

## Exploitation des données en temps réel : beaucoup reste à faire pour relever le défi

Analyste : Thierry Hamelin  
Janvier 2022

Sponsorisé par Confluent

# INTRODUCTION

---

Le contexte semble favorable au développement des stratégies data-driven. En effet, les organisations accordent une importance croissante à la valorisation de leurs données et 74% d'entre-elles considèrent que disposer d'informations en temps réel représentera un atout pour leur activité dans les années à venir.

Toutefois, 93% des organisations sont confrontées à des difficultés pour exploiter leurs données en temps réel. C'est ce que révèle l'étude menée par IDC en France auprès de 200 organisations privées et publiques, dont les résultats sont présentés dans ce document.

Une majorité d'organisations rencontre des problèmes relatifs à la capacité de leurs applications à gérer le temps réel, à l'intégration de multiples sources de données ou encore au volume de données à traiter.

Or la situation va aller en empirant si rien n'est fait. En effet, les organisations étudiées anticipent une augmentation significative du volume de données qu'elles vont avoir à gérer dans les cinq années à venir et dans le même temps, un renforcement de leurs besoins pour disposer d'informations en temps réel.

A cela s'ajoute l'intégration de nouvelles sources de données, en particulier celles issues de leurs partenaires commerciaux et de leurs fournisseurs dans un souci d'efficacité et de résilience ou encore celles résultant de leurs interactions avec les clients, voire les administrés ou les patients selon le secteur d'activité, en vue d'améliorer l'expérience utilisateur.

Dans ce contexte, Apache Kafka, une plate-forme open source de streaming d'événements développée par Confluent et conçue pour traiter en temps réel un flux massif de données pour les stocker et les rendre accessibles à toutes les applications de l'entreprise suscite l'intérêt. En effet, 51% des organisations étudiées considèrent cette plate-forme comme intéressante.

C'est son large éventail d'usages, comme le pilotage de la production, le décisionnel, la cybersécurité ou la gestion de la relation client, qui est plébiscité. Autant de domaines dont l'étude d'IDC montre que les besoins pour disposer d'informations en temps réel vont se renforcer dans les années à venir.

# PRÉSENTATION DE L'ÉTUDE

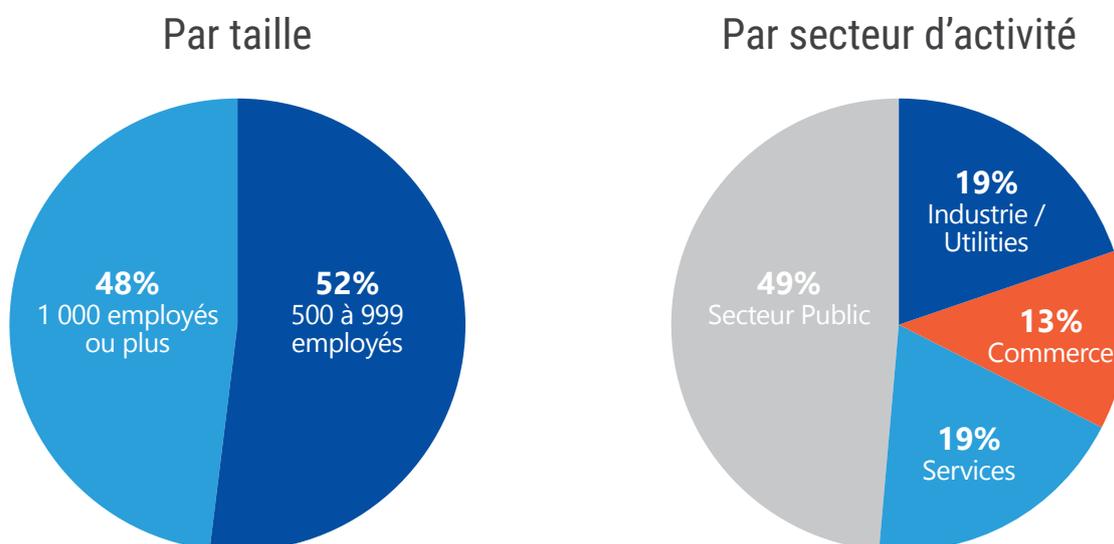
Les résultats présentés dans ce document reposent sur une enquête téléphonique menée en France par IDC entre juillet et septembre 2021 auprès de 200 organisations comptant au moins 500 employés. L'échantillon étudié se compose de 51% d'organisations du secteur privé (industrie, services et commerce) et 49% du secteur public (collectivités territoriales et hôpitaux).

Les personnes interrogées sont des décideurs IT (DSI et responsables informatiques).

Les résultats ont été redressés conformément aux statistiques de l'Insee, en termes d'effectif salarié et de secteur d'activité.

## GRAPHIQUE 1

Profil des 200 organisations étudiées



Source : IDC France, étude gestion des données 2021, n=200



# UNE IMPORTANCE CROISSANTE ACCORDÉE À LA VALORISATION DES DONNÉES

## D'ici trois ans,

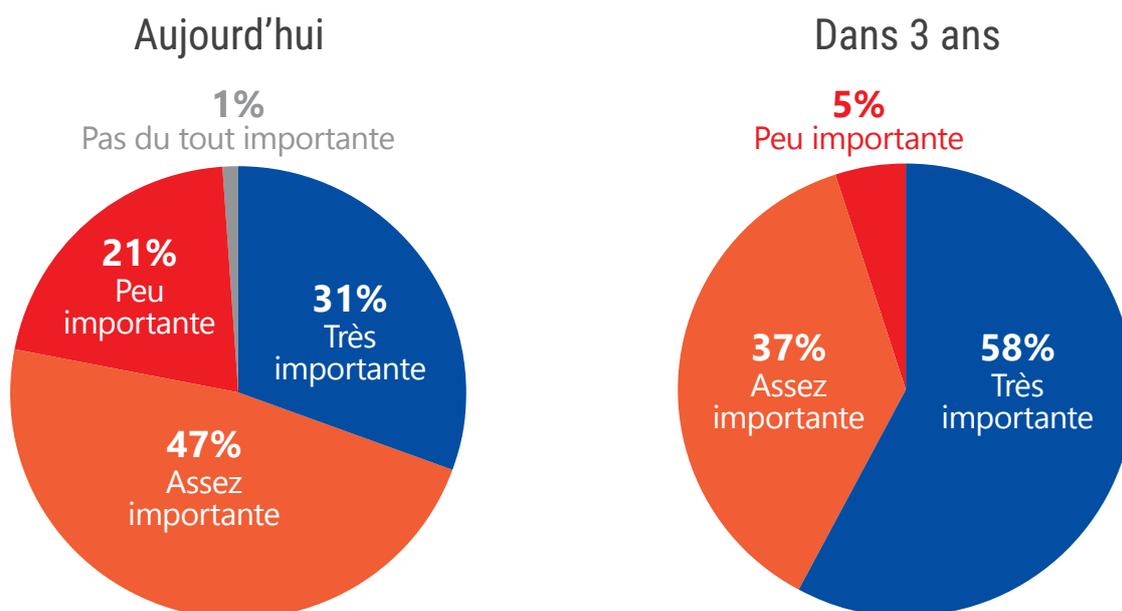
58% des organisations estiment que la valorisation des données sera "très importante" pour elles, contre 31% actuellement.

Même si elles y sont déjà sensibles, les organisations vont accorder une importance croissante à la valorisation de leurs données dans les années à venir, comme en atteste le graphique 2. D'ici trois ans, 58% estiment que la valorisation des données sera "très importante" pour elles, contre 31% actuellement. Par ailleurs, si 21% la considèrent comme "peu importante" à ce jour, elles ne seront plus que 5% dans trois ans.

### GRAPHIQUE 2

Importance accordée à la valorisation des données

Q. Quelle importance votre organisation accorde-t-elle à la valorisation des données et qu'en sera-t-il d'ici 3 ans ?



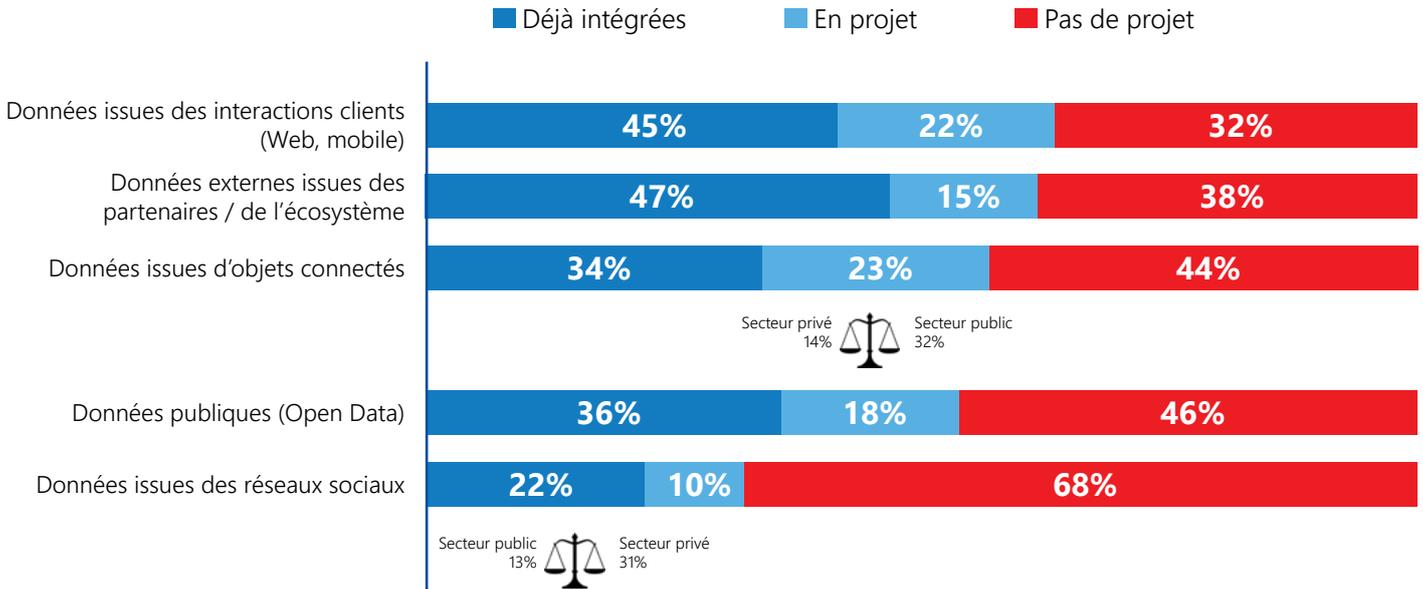
Source : IDC France, étude gestion des données 2021, n=200

Le graphique 3 montre que les organisations ont également engagé des stratégies visant à intégrer de nouvelles sources de données. C'est notamment le cas des données liées aux interactions avec les clients ou encore les administrés et patients selon le secteur d'activité. Cela va donc nécessiter de pouvoir gérer des données non-structurées, à l'instar des informations issues des réseaux sociaux, même si c'est pour l'instant un domaine qui se révèle moins prioritaire.

### GRAPHIQUE 3

#### Exploitation de nouvelles sources de données

#### Q. Quelles nouvelles sources de données intégrez-vous ou prévoyez-vous d'intégrer dans vos projets de valorisation de données ?



Source : IDC France, étude gestion des données 2021, n=200

Les organisations cherchent également à intégrer des données issues de leur écosystème de partenaires commerciaux et de fournisseurs pour gagner en efficacité et en visibilité sur leur activité.

Enfin, les objets connectés et les informations publiques (open data) sont également des domaines où les organisations cherchent à valoriser leurs données. Le secteur public se montre sensible aux objets connectés puisque 32% des organisations sont engagées dans des projets.



# UNE FORTE AUGMENTATION ATTENDUE DU VOLUME DE DONNÉES

## Dans le secteur privé,

48% des entreprises estiment que le décisionnel va contribuer à fortement augmenter le volume de données dans les 5 ans à venir.

Quel que soit leur secteur d'activité, les organisations étudiées anticipent toutes une augmentation significative du volume de données qu'elles vont être amenées à gérer dans les années à venir.

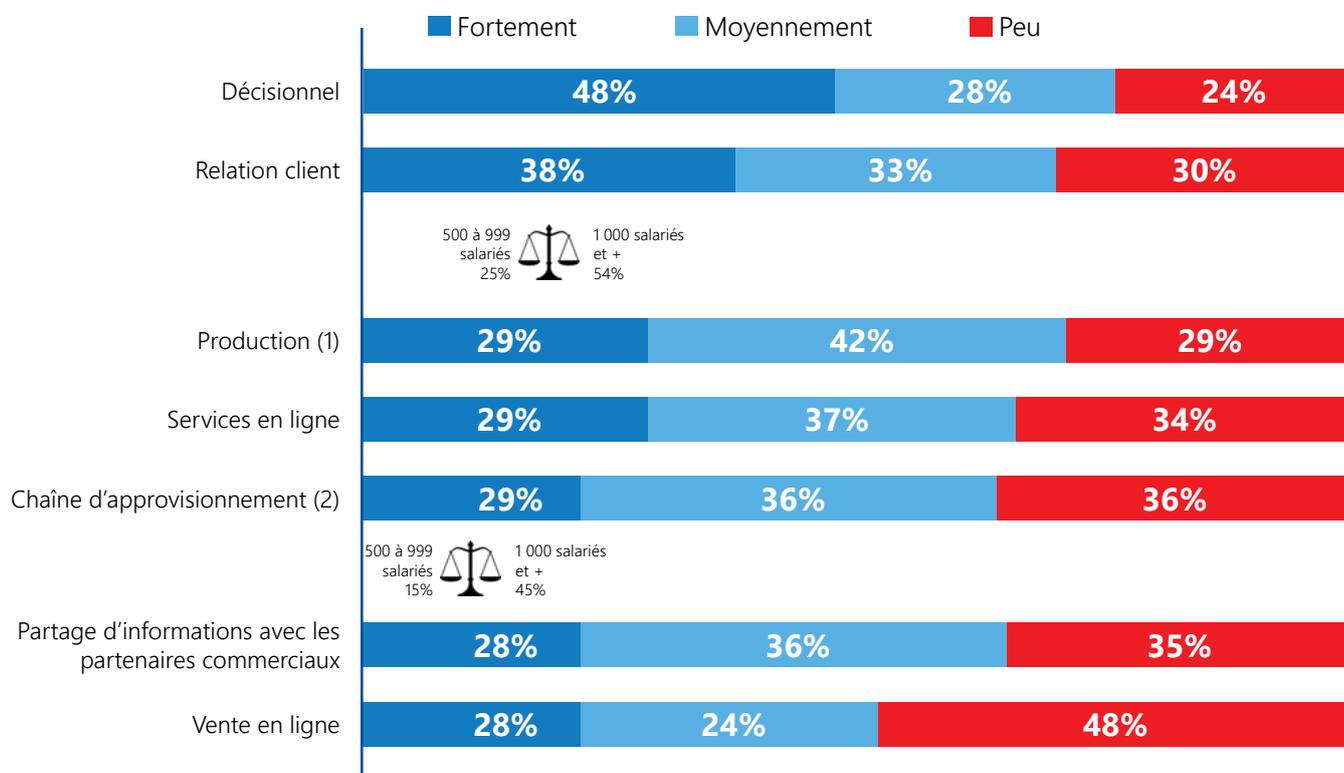
Dans le secteur privé, de nombreux éléments vont y contribuer, comme le montre le graphique 4. C'est notamment le cas du décisionnel pour lequel près d'une entreprise sur deux table sur une forte augmentation du volume de données généré.

La relation client et la gestion de la chaîne d'approvisionnement sont des domaines où les grandes entreprises sont bien plus nombreuses à anticiper une forte hausse du volume de données qu'elles auront à traiter dans les 5 ans à venir. Un constat qui découle de la nécessité d'optimiser les interactions avec les clients, quel que soit le canal, et d'améliorer la résilience en anticipant le plus possible les ruptures d'approvisionnement. Une approche qui va de plus en plus gagner les plus petites organisations à l'avenir.

### GRAPHIQUE 4

#### Augmentation du volume de données (secteur privé)

**Q. Selon vous, dans les 5 prochaines années, dans quelle mesure votre processus de transformation numérique va-t-il contribuer à augmenter le volume de données à gérer pour votre organisation dans les domaines suivants ?**



Source : IDC France, étude gestion des données 2021, n=103  
 (1) Entreprises du secteur de l'industrie (n=40)  
 (2) Entreprises du secteur de l'industrie (n=40) et du commerce (n=25)

# Dans les collectivités territoriales,

les plus fortes hausses attendues du volume de données seront liées aux services internes, à la relation avec les administrés ainsi que les services en ligne qui leur sont destinés.

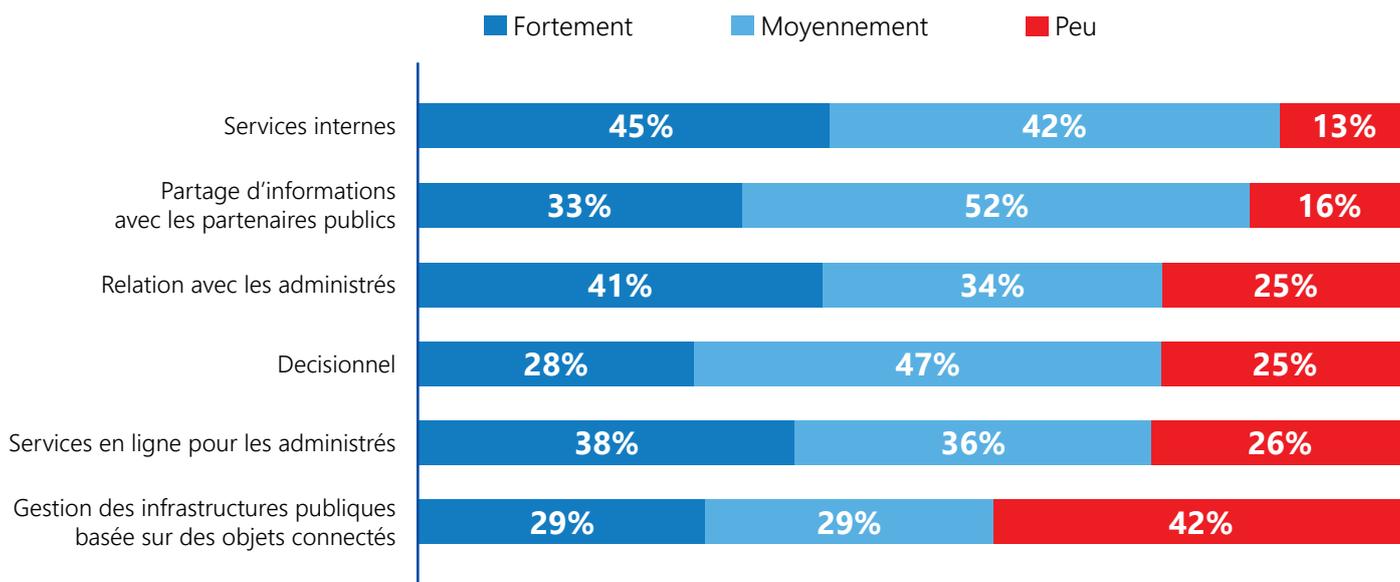
Dans les collectivités territoriales, la situation est comparable puisque de nombreux domaines vont contribuer à augmenter le volume de données à gérer. Comme le prouve le graphique 5, les plus grosses hausses sont attendues dans les services internes, la relation avec les administrés et les services en ligne qui leur sont destinés.

Les objets connectés arrivent en retrait car leur utilisation est encore loin d'être généralisée au sein des collectivités territoriales, et celles qui n'y recourent pas encore ont du mal à anticiper leur impact sur la croissance du volume de données. Ainsi, 20% des collectivités n'utilisant pas les objets connectés estiment que cela va "fortement" contribuer à augmenter leur volume de données, alors que cette part atteint 48% chez celles qui y recourent déjà.

## GRAPHIQUE 5

### Augmentation du volume de données (collectivités territoriales)

Q. Selon vous, dans les 5 prochaines années, dans quelle mesure votre processus de transformation numérique va-t-il contribuer à augmenter le volume de données à gérer pour votre organisation dans les domaines suivants ?



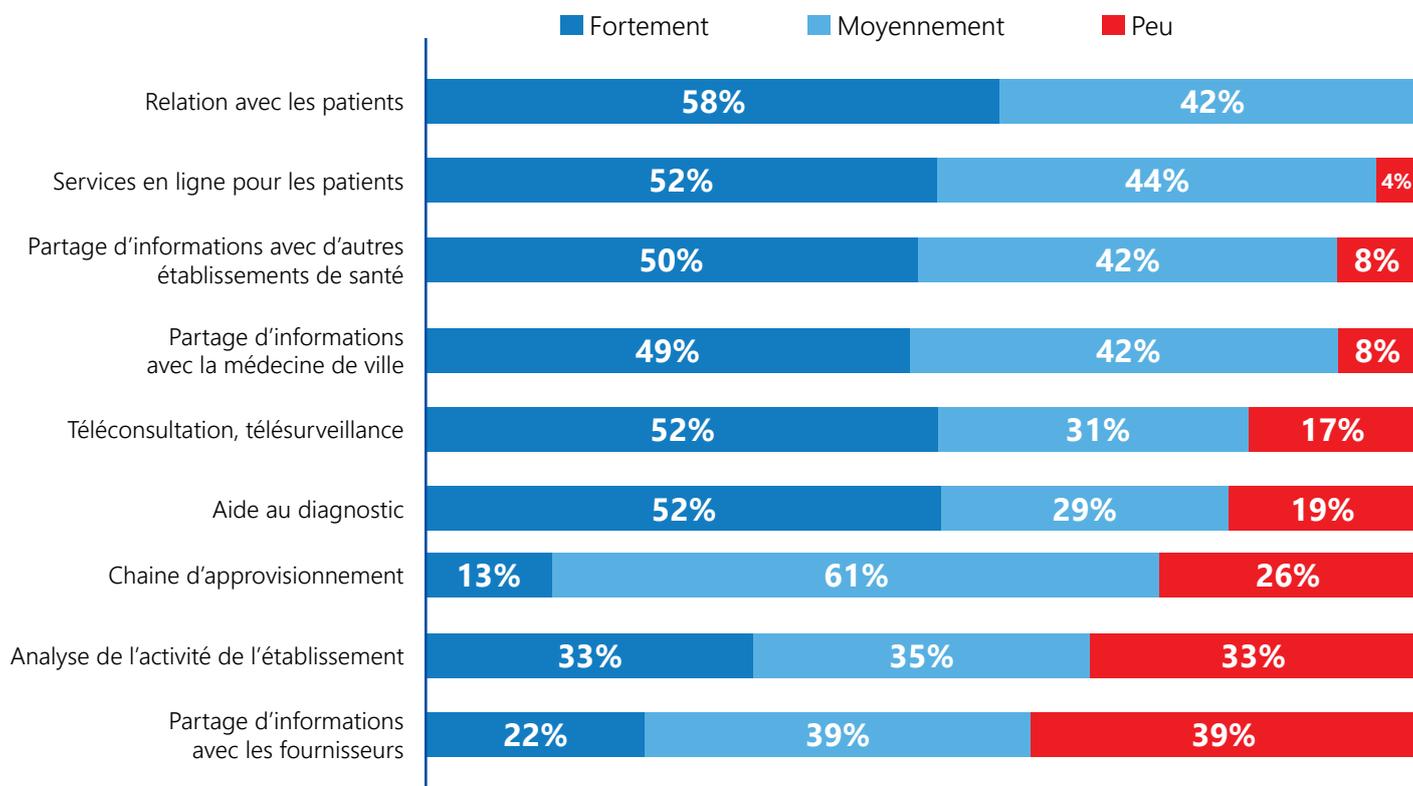
Source : IDC France, étude gestion des données 2021, n=73

C'est pour les hôpitaux que la croissance anticipée du volume de données est la plus forte. Comme le montre le graphique 6, la moitié environ table sur le fait que de nombreux domaines vont conduire à fortement augmenter le volume de données à gérer dans les 5 prochaines années. C'est le cas de la relation avec les patients, les services en ligne destinés aux patients, le partage d'informations avec d'autres établissements de soins et avec la médecine de ville, la téléconsultation / la télésurveillance et enfin l'aide au diagnostic.

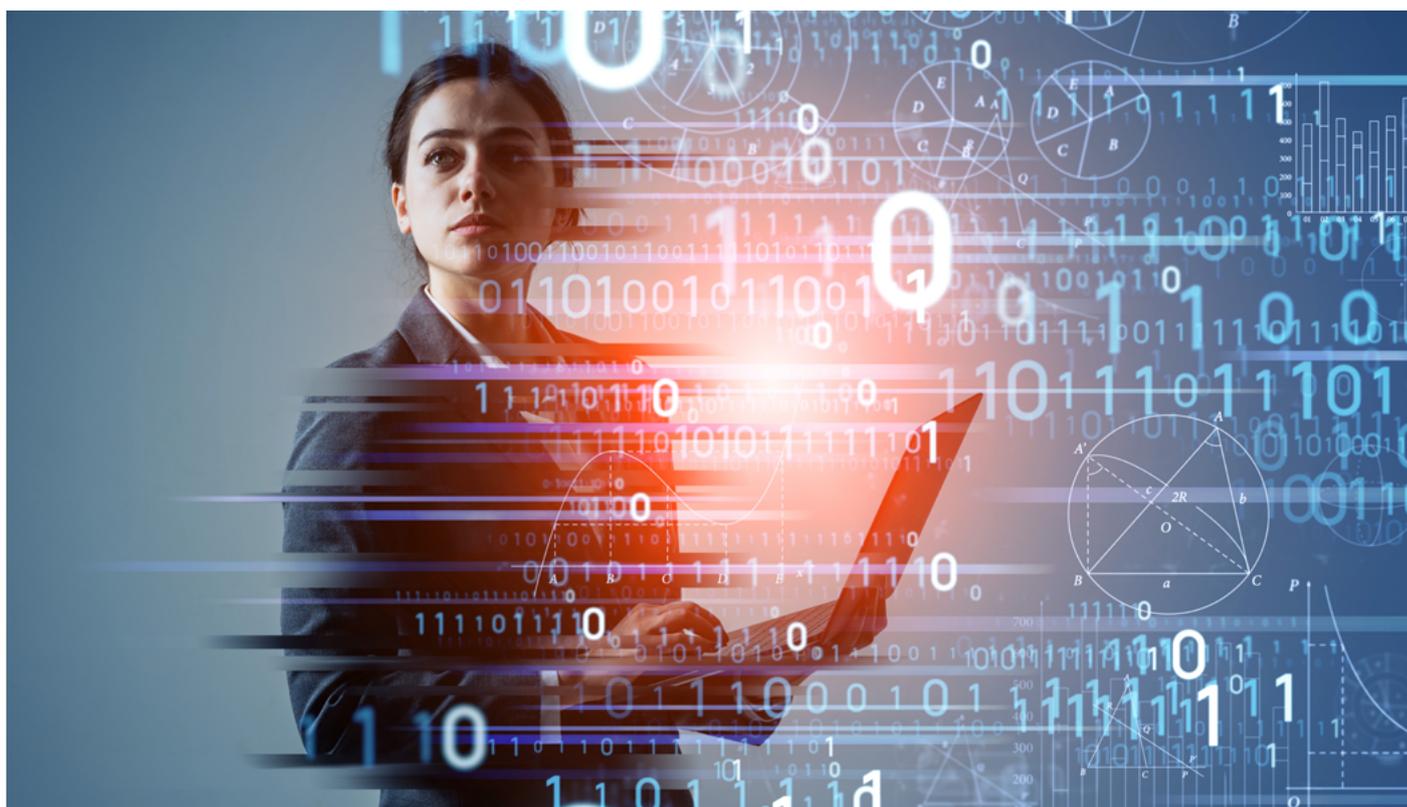
## GRAPHIQUE 6

### Augmentation du volume de données (hôpitaux)

**Q. Selon vous, dans les 5 prochaines années, dans quelle mesure votre processus de transformation numérique va-t-il contribuer à augmenter le volume de données à gérer pour votre organisation dans les domaines suivants ?**



Source : IDC France, étude gestion des données 2021, n=24



# VERS UN RENFORCEMENT DES FLUX DE DONNÉES EN TEMPS RÉEL

Les organisations, quel que soit leur secteur d'activité, estiment que c'est pour gérer leur IT, et surtout la cybersécurité, que leurs besoins de disposer de données en temps réel va le plus se renforcer dans les prochaines années. Au-delà de ces usages, on constate logiquement des différences selon leur domaine d'activité.

## Dans le secteur privé, les besoins pour davantage de données en temps réel concernent l'IT mais également la production et les clients

### Dans le secteur privé,

51% des entreprises anticipent que la gestion de l'IT sera une source croissante de données en temps réel, contre 40% pour la production et 36% pour la relation client.

Dans le secteur privé, les données pour gérer l'IT et la cybersécurité ainsi que la production sont d'ores et déjà celles qui sont le plus souvent gérées en temps réel. Ce constat va encore se renforcer à l'avenir, comme en atteste le graphique 7. Avec la montée en puissance de l'industrie 4.0, le pilotage de la production basé sur une utilisation plus forte de l'informatique et l'intégration de nouvelles technologies comme les objets connectés va renforcer les flux de données temps réel.

Les informations relatives à la relation client seront également une source croissante de données en temps réel pour les entreprises dans les cinq prochaines années. L'enjeu est d'améliorer les échanges avec les clients, quel que soit le canal, grâce à des interactions personnalisées et contextualisées, mais aussi de mieux gérer les réclamations. La réactivité est un facteur fondamental en la matière qui dépend de la capacité des entreprises à gérer des informations en temps réel.

Il en va de même pour les services en ligne permettant aux clients de réaliser eux-mêmes certaines opérations mais aussi pour les informations relatives aux commandes qui doivent être traitées au plus vite, au même titre que celles liées à la chaîne d'approvisionnement pour prévenir toute rupture dans le processus de production.

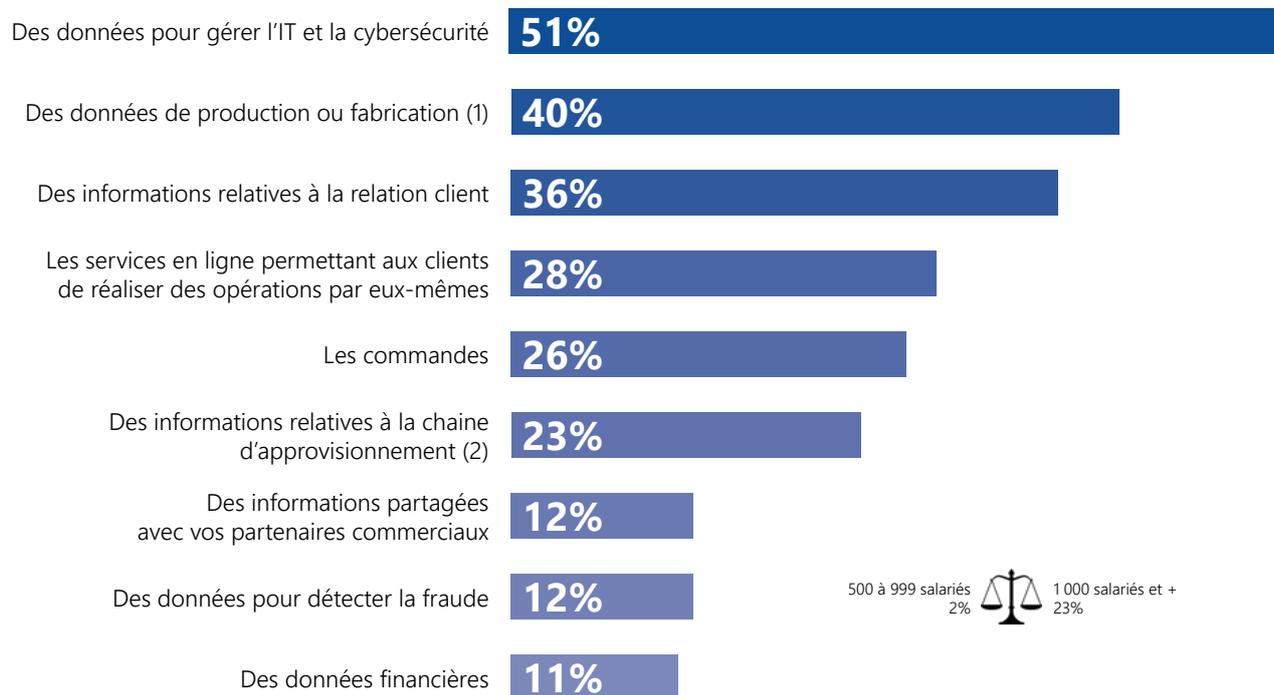
Enfin, près d'un quart des grandes entreprises mise sur une utilisation croissante de données en temps réel pour lutter contre la fraude, alors que cela ne concerne pour l'instant que 2% des plus petites organisations.



## GRAPHIQUE 7

### Besoins pour des flux de données en temps réel (secteur privé)

Q. Quels sont les flux de données pour lesquels vous anticipez un besoin croissant de votre organisation pour disposer de données en temps réel dans les 5 prochaines années ?



Source : IDC France, étude gestion des données 2021, n=103  
(1) Entreprises du secteur de l'industrie (n=40)  
(2) Entreprises du secteur de l'industrie (n=40) et du commerce (n=25)

## 67% des collectivités territoriales

estiment que la gestion de l'IT sera une source croissante de données en temps réel, contre 40% pour la gestion des infrastructures publiques basée sur les objets connectés.

### Dans les collectivités territoriales, au-delà de l'IT, les besoins pour plus de données temps réel résultent de la gestion des infrastructures publiques, des services internes et des administrés

Dans les collectivités territoriales, à l'instar du secteur privé, la gestion de l'IT et de la cybersécurité sont actuellement la première source de données temps réel. Le graphique 8 confirme que cette situation va perdurer à l'avenir mais révèle que d'autres domaines vont contribuer à renforcer les flux de données en temps réel.

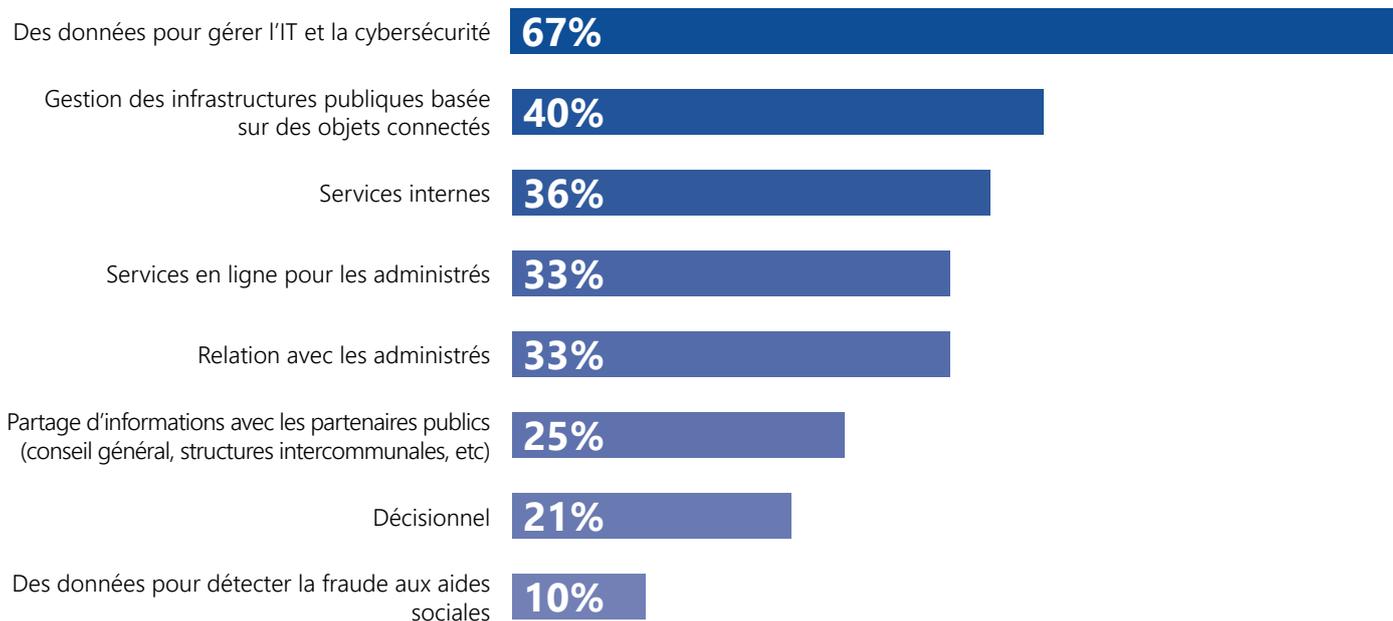
C'est notamment le cas de la gestion des infrastructures publiques qui, avec l'essor des objets connectés, va de plus en plus s'appuyer sur des données temps réel pour mieux réguler le trafic, les risques environnementaux, la sécurité, etc.

Il en va de même des domaines pour lesquels la réactivité est un élément essentiel comme les services internes, les services en ligne et les relations avec les administrés.

## GRAPHIQUE 8

### Besoins pour des flux de données en temps réel (collectivités territoriales)

Q. Quels sont les flux de données pour lesquels vous anticipez un besoin croissant de votre organisation pour disposer de données en temps réel dans les 5 prochaines années ?



Source : IDC France, étude gestion des données 2021, n=73

### Dans les hôpitaux, les besoins pour davantage de données en temps réel pour la gestion des patients talonnent la gestion de l'IT

**61% des hôpitaux** anticipent que la gestion de l'IT sera une source croissante de données en temps réel, contre 57% pour la gestion de la relation avec les patients.

Dans les hôpitaux publics, comme pour les autres secteurs d'activité, les données pour gérer l'IT et la cybersécurité arrivent déjà en tête des flux gérés en temps réel et cela va se poursuivre dans les cinq ans à venir comme le montre le graphique 9. Ce dernier révèle aussi que de nombreux domaines (les mêmes que ceux qui vont générer les plus fortes croissances du volume de données) vont conduire à une augmentation des besoins pour des flux en temps réel.

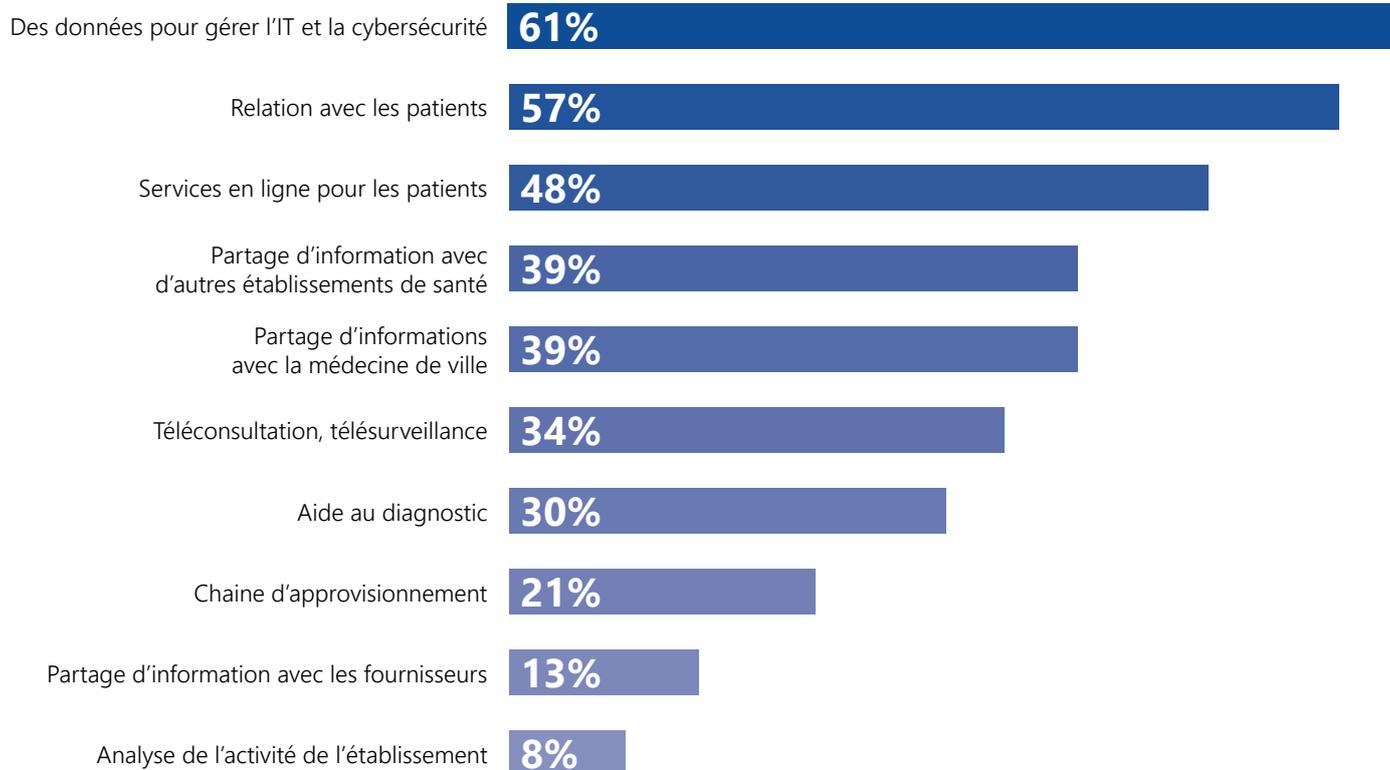
C'est notamment le cas de la relation avec les patients et des services en ligne qui leur sont destinés, et dans une moindre mesure, du partage d'informations avec d'autres établissements de soins ou la médecine de ville, de la téléconsultation / télésurveillance ou encore de l'aide au diagnostic.

Le programme "Hop'En" (Hôpital numérique ouvert sur son environnement) engagé par l'Etat joue un rôle moteur en la matière puisqu'il vise à développer les services numériques à destination des patients et l'ouverture des systèmes d'information des hôpitaux à la médecine de ville, que ce soit pour le partage d'avis médicaux, l'imagerie, les agendas partagés, etc.

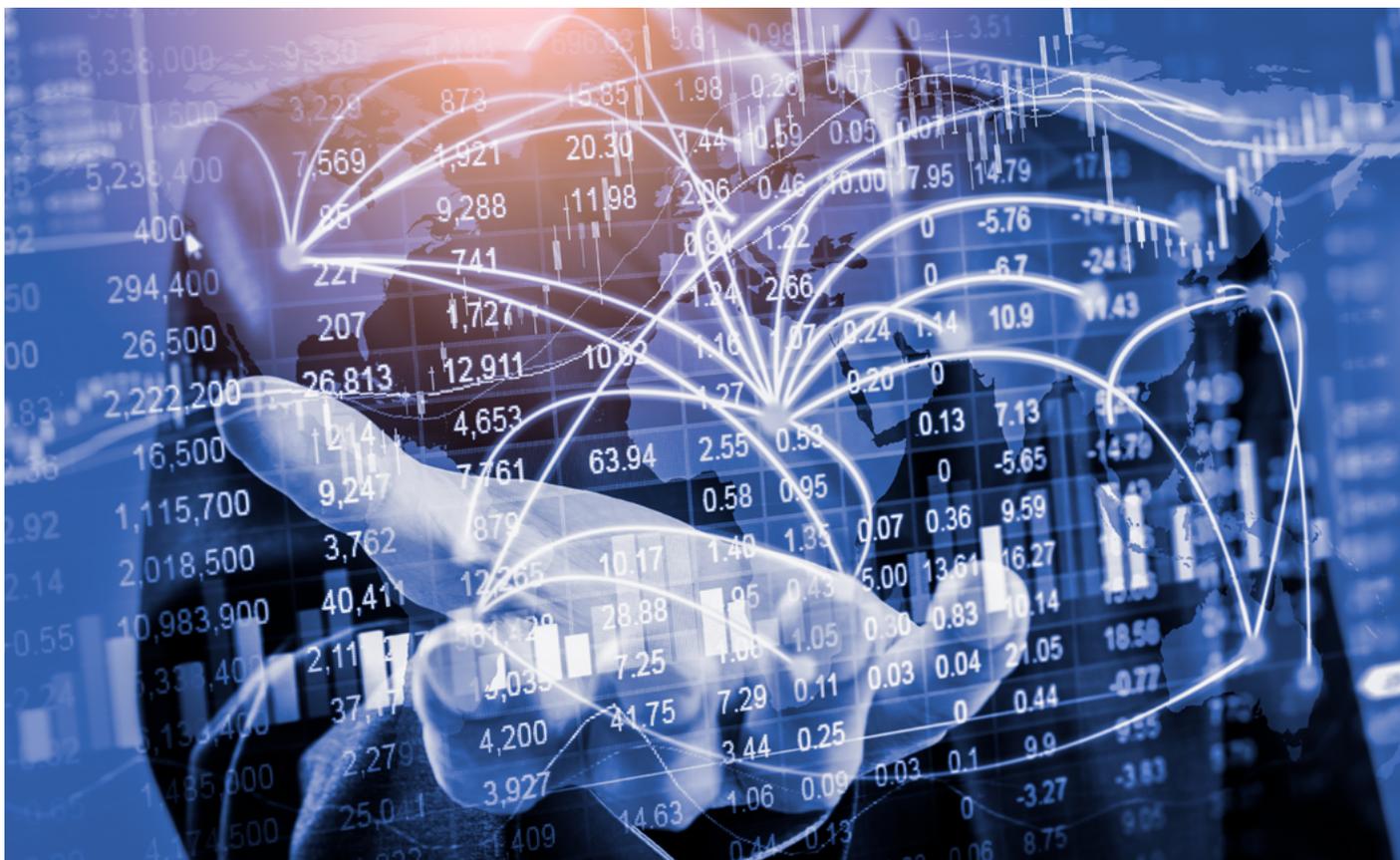
## GRAPHIQUE 9

### Besoins pour des flux de données en temps réel (hôpitaux)

**Q. Quels sont les flux de données pour lesquels vous anticipez un besoin croissant de votre organisation pour disposer de données en temps réel dans les 5 prochaines années ?**



Source : IDC France, étude gestion des données 2021, n=24



# DES DIFFICULTÉS PESANT SUR L'EXPLOITATION DES DONNÉES EN TEMPS RÉEL

## Plus de deux tiers

des organisations rencontrent des difficultés pour exploiter leurs données en temps réel du fait de l'incompatibilité de leurs applications métiers ou de la multiplicité des sources d'information.

Le contexte semble favorable au développement de stratégies data-driven au sein des organisations. En effet, elles accordent une importance croissante à la valorisation de leurs données et 74% d'entre-elles considèrent que disposer de données en temps réel représentera un atout pour leur activité dans les 5 ans à venir. Cependant, 93% des organisations que nous avons étudiées sont confrontées à au moins une difficulté pour exploiter leurs données en temps réel.

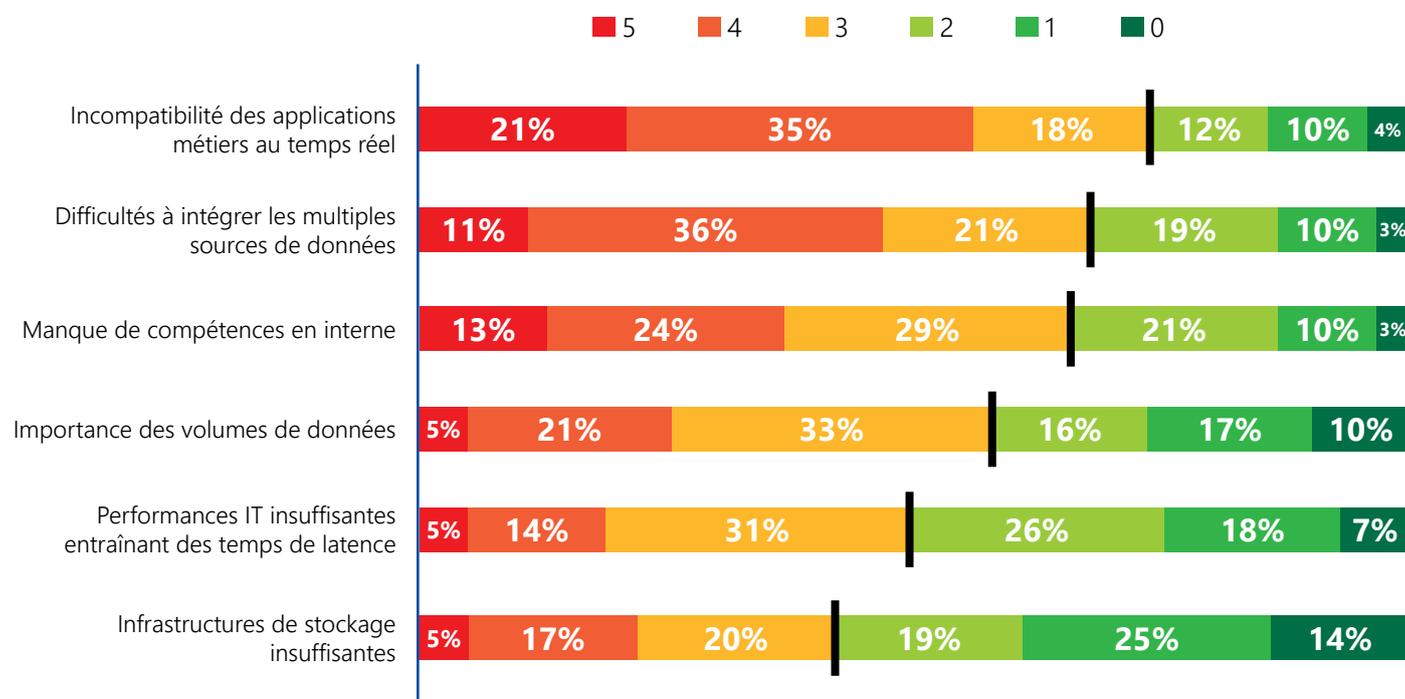
Comme le montre le graphique 10, une majorité d'organisations rencontre des problèmes relatifs à la capacité de leurs applications à gérer le temps réel, à l'intégration de multiples sources de données, au déficit de compétences en interne ou encore au volume de données à traiter.

Par ailleurs, la moitié des organisations fait face à des temps de latence liés à la performance insuffisante de leur IT. Si les capacités de stockage semblent pour l'instant moins problématiques, la croissance du volume de données à laquelle les organisations vont être confrontées dans les années à venir pourrait venir relativiser ce constat.

### GRAPHIQUE 10

#### Difficultés rencontrées en matière d'exploitation des données en temps réel

Q. Sur une échelle de 0 (pas du tout important) à 5 (très important) à quel point les facteurs suivants vous posent-ils des difficultés dans l'exploitation des données en temps réel ?



Source : IDC France, étude gestion des données 2021, n=200

S'agissant plus particulièrement des difficultés rencontrées en matière d'intégration de multiples sources de données, le graphique 11 montre qu'elles pèsent sur la performance des organisations.

## Le décisionnel est le domaine

où le manque d'intégration entre les différentes sources de données pèse le plus sur la performance.

C'est surtout le cas pour le décisionnel, qui combine diverses sources de données souvent cloisonnées, et dans une moindre mesure pour les relations avec les clients ou encore les administrés et patients afin de bénéficier d'une vision 360 degrés, quel que soit le canal d'interaction.

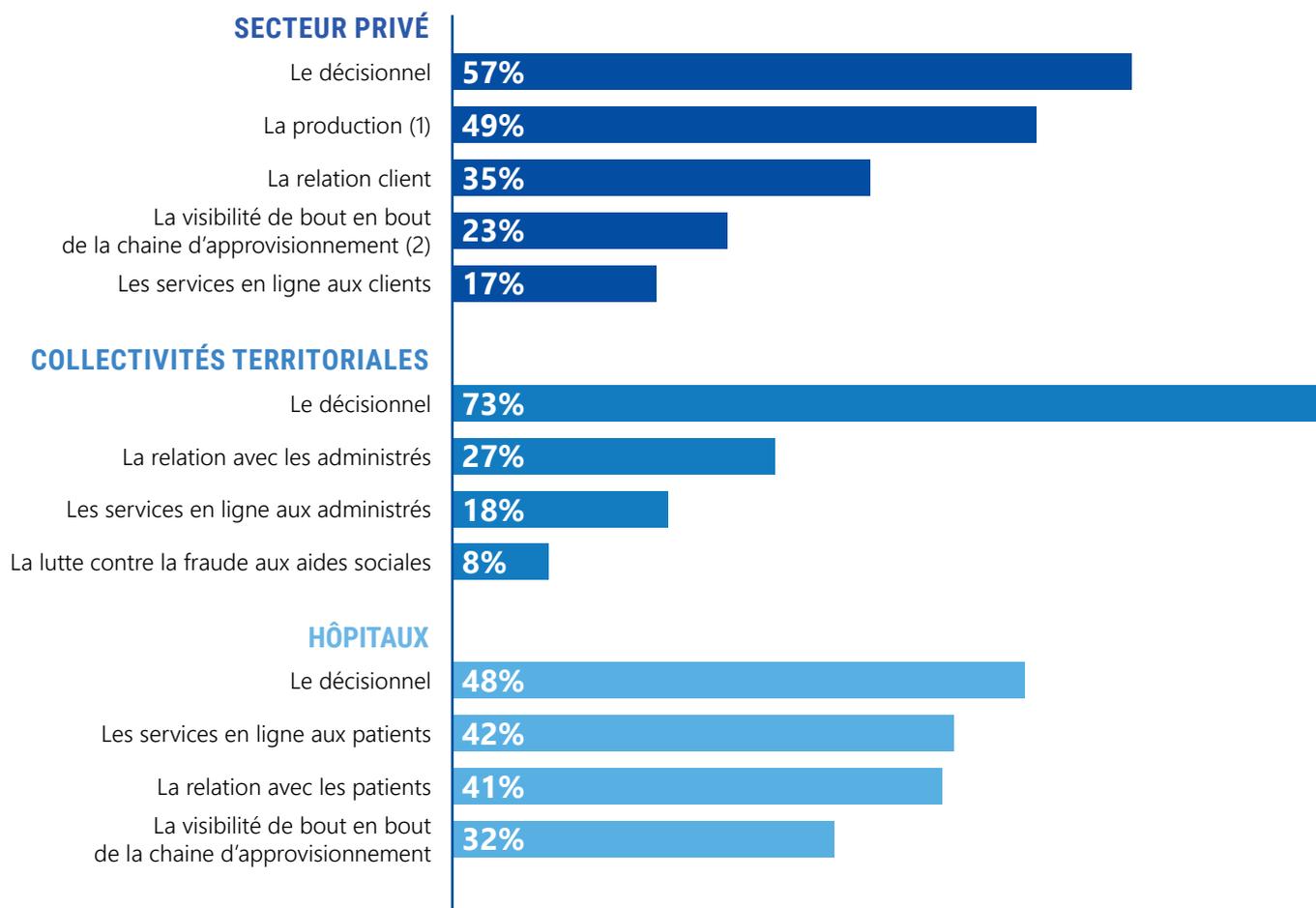
Au sein du secteur privé, les entreprises industrielles se distinguent par les difficultés qu'elles rencontrent pour ajuster la production en fonction de données issues de multiples sources, comme les informations commerciales ou encore celles liées aux approvisionnements. Elles sont 49% à souffrir d'un problème d'intégration qui impacte leur performance.

Enfin, au sein des secteurs du commerce et de l'industrie, 23% des entreprises rencontrent des difficultés pour disposer d'une visibilité de bout en bout sur leur chaîne d'approvisionnement, tout comme les hôpitaux qui sont près d'un tiers dans cette situation.

### GRAPHIQUE 11

#### Niveau d'avancement de la stratégie Industrie 4.0

Q. Quels sont les domaines où le manque d'intégration entre les différentes sources de données pèse sur la performance ?



Source : IDC France, étude gestion des données 2021  
 Secteur privé (n= 103) / Collectivités territoriales (n=73) / Hôpitaux (n=24)  
 (1) Entreprises du secteur de l'industrie (n=40)  
 (2) Entreprises du secteur de l'industrie (n=40) et du commerce (n=25)

## Un traitement des données qui reste très perfectible

### 29% des organisations

ne sont pas encore équipées d'un entrepôt de données et cette part atteint 39% chez celles ayant moins de 1 000 salariés.

29% des organisations étudiées ne sont pas encore équipées d'un entrepôt de données et cette proportion atteint 39% chez celles ayant moins de 1 000 salariés. Or le graphique 12 révèle que seules 24% des organisations qui ne disposent pas encore d'un entrepôt de données comptent franchir le pas sous 24 mois quand dans le même temps, 64% restent sans aucun projet d'équipement.

Parmi les organisations équipées, près de deux tiers disposent de plusieurs entrepôts de données, ce qui peut compliquer le traitement de données issues de multiples sources. Par ailleurs, 70% des organisations s'appuient encore sur une solution on-premise susceptible de brider les capacités de montée en puissance.

Enfin, un quart des organisations s'appuie sur des équipements ayant plus de cinq ans et un tiers sur des solutions ayant de deux à cinq ans. Or le graphique 12 montre que seules 32% des organisations envisagent de moderniser leur solution sous 24 mois.

### GRAPHIQUE 12

#### Modernisation des entrepôts de données

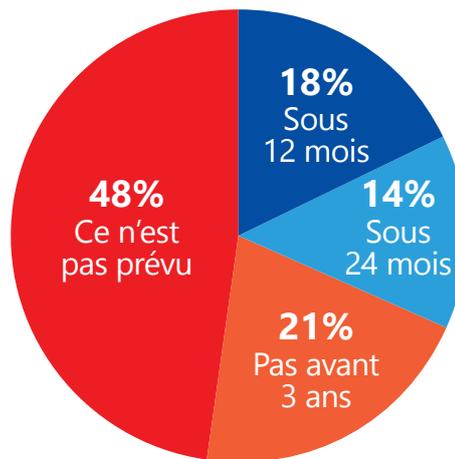
Q. Quand envisagez-vous de vous équiper d'un entrepôt de données ?



N=58

(organisations n'utilisant pas encore d'entrepôt de données)

Q. Quand envisagez-vous de moderniser votre équipement ?



N=142

(organisations utilisant au moins un entrepôt de données)

De son côté, le Big Data est encore loin d'être généralisé puisque seule une organisation sur cinq y recourt. Cette part chute d'ailleurs à 11% dans le secteur public. Si 13% des organisations étudient ou prévoient un équipement sous 24 mois, 67% restent sans projet.

Malgré la volonté de mieux valoriser leurs données, d'évoluer vers des traitements en temps réel et de mieux gérer leur relation client / administré / patient, les organisations se révèlent assez mal armées.

En effet, elles s'appuient encore en majorité sur des entrepôts de données nécessitant une réplique des données, ce qui pèse sur leur capacité à exploiter les informations en temps réel. De plus, avec l'essor des données non-structurées qui caractérisent notamment la relation client, administré, patient, les entrepôts de données se révèlent assez inappropriés.

# APACHE KAFKA SUSCITE L'INTÉRÊT

**51% des organisations** considèrent la plate-forme Apache Kafka comme intéressante.

Apache Kafka est une plate-forme open source de streaming d'événements conçue pour traiter en temps réel un flux massif de données pour les stocker et les rendre accessibles à toutes les applications de l'entreprise. Ce sont les fondateurs de la société Confluent qui sont à l'initiative de Kafka. Même si la plate-forme est désormais en open source, 85% du code de Kafka est encore développé par Confluent.

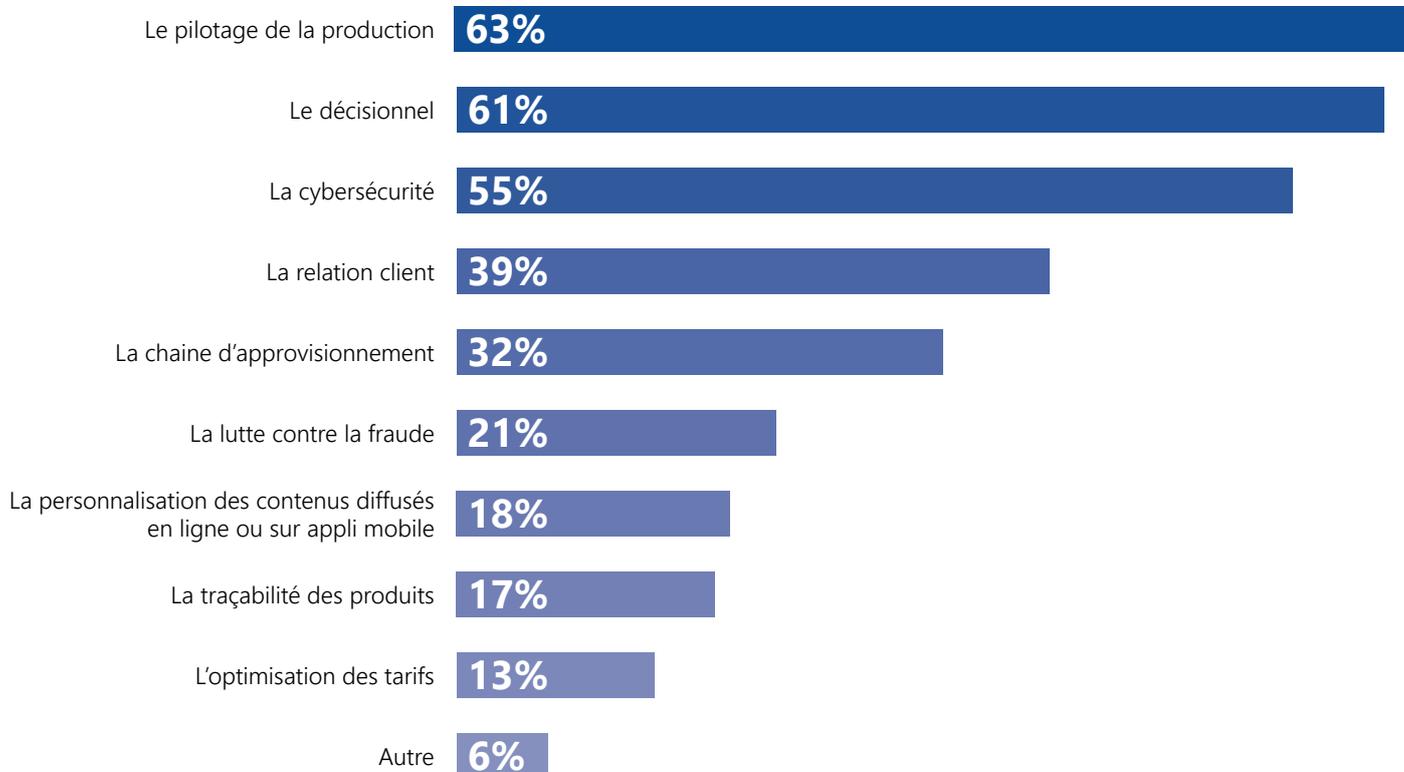
Il s'avère que 51% des organisations étudiées considèrent la plate-forme Apache Kafka comme intéressante. D'ailleurs, 32% en avaient déjà entendu parler.

Apache Kafka suscite l'intérêt par son large éventail d'usages. C'est particulièrement le cas pour le pilotage de la production, le décisionnel et la cybersécurité, et dans une moindre mesure, de la relation client au sens large (c'est-à-dire incluant les administrés et les patients). Autant de domaines dont on a vu que les besoins pour disposer d'informations en temps réel vont se renforcer dans les années à venir.

## GRAPHIQUE 13

### Intérêt pour Apache Kafka

**Q. Pour quels usages trouvez-vous cette plate-forme de streaming d'événements particulièrement intéressante ?**



Source : IDC France, étude gestion des données 2021, n=102 (organisations trouvant intéressante la plateforme Apache Kafka)

# L'ÉCLAIRAGE DE CONFLUENT

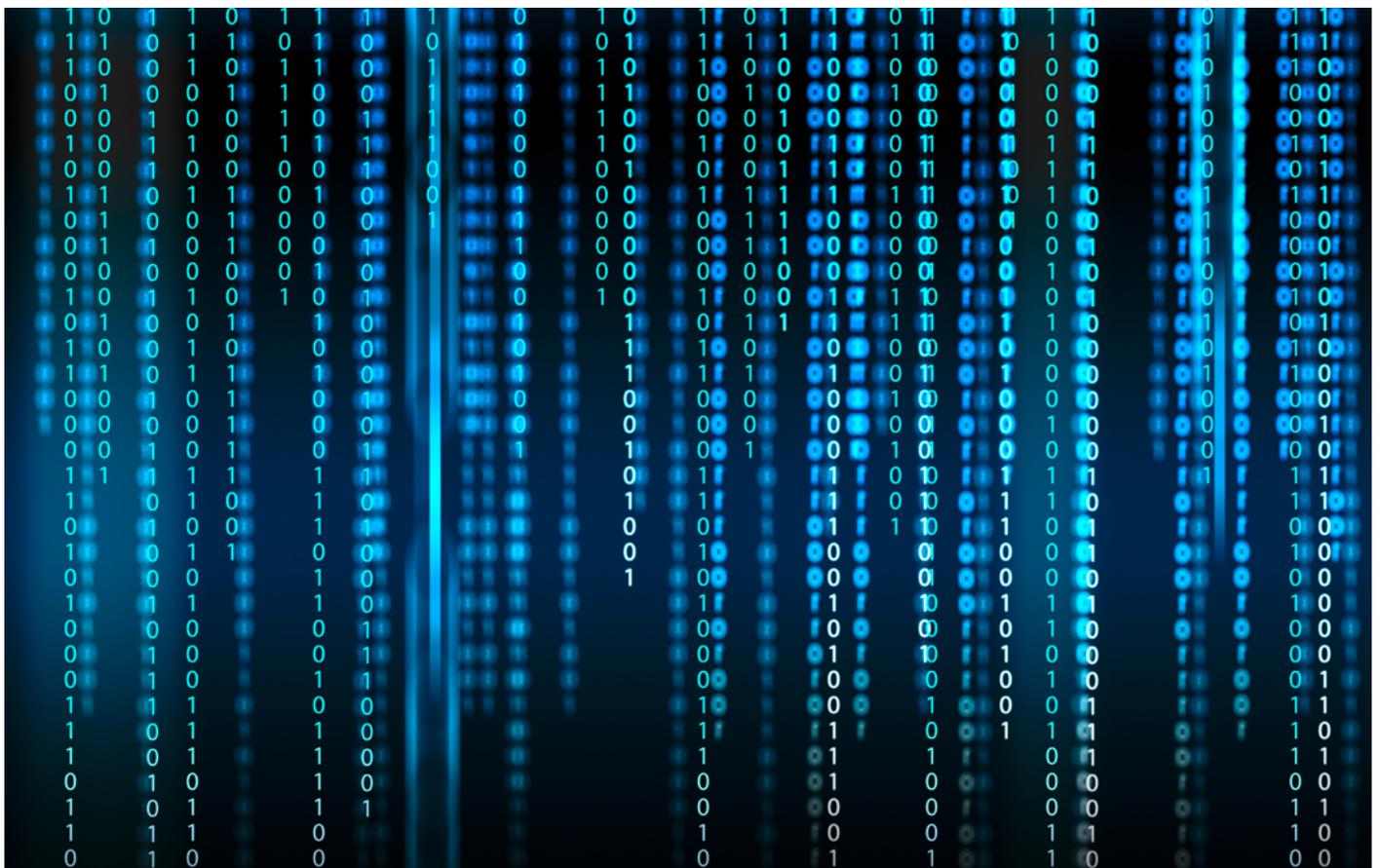
Confluent est le pionnier d'un nouveau type d'infrastructure de données que nous appelons data in motion. Au cœur de ce mouvement se trouve la technologie open-source Apache Kafka®, que plus de 70 % des entreprises du Fortune 500 utilisent aujourd'hui.

Alors que les entreprises passent à des approches numériques, de grandes parties de leurs organisations dépendent des logiciels. Cela ne se limite pas uniquement aux interactions des clients sur le web, mais cela touche également la logistique, les processus de fabrication et toutes les étapes complexes pour coordonner les biens et les services. C'est pourquoi la data in motion, ou données en temps réel, est vraiment essentielle.

L'étude menée par IDC fournit un retour révélateur sur l'importance croissante des données en temps réel sur le marché français, avec 74% des répondants qui considèrent que ce sera un élément important pour leur activité dans les années à venir. Cependant, l'enquête montre que 93% des entreprises font face à des défis importants pour tirer parti de leurs données en temps réel.

D'autres résultats sont intéressants. Ainsi, 58% des répondants considèrent que tirer parti des données sera "très important" pour leur entreprise d'ici 3 ans, contre 31% actuellement. A l'inverse, 21% estiment que leur entreprise y accorde "peu importance" à ce jour mais seuls 5% anticipent que ce sera encore le cas dans 3 ans.

Au vu des résultats de cette étude, nous sommes encore plus enthousiastes d'aider les entreprises à surmonter leurs défis et mettre leur data in motion pour offrir d'excellentes expériences à leurs employés, leurs partenaires commerciaux et à leurs clients.



## À PROPOS D'IDC

IDC est un acteur majeur de la Recherche, du Conseil et de l'Évènementiel sur les marchés des Technologies de l'Information, des Télécommunications et des Technologies Grand Public. IDC aide les professionnels évoluant sur les marchés IT et les investisseurs à prendre des décisions stratégiques basées sur des données factuelles. Plus de 1 100 analystes proposent leur expertise globale, régionale et locale sur les opportunités et les tendances technologies dans plus de 110 pays à travers le monde. Depuis plus de 50 ans, IDC propose des analyses stratégiques pour aider ses clients à atteindre leurs objectifs clés. IDC est une filiale de la société IDG, leader mondial du marché de l'information dédiée aux technologies de l'information.

### IDC FRANCE

102-116 Rue Victor Hugo

92300 Levallois-Perret

+33 1 56 26 26 66

Twitter : @IDCFrance

[idc-community.com](http://idc-community.com)

[www.idc.com](http://www.idc.com) / [www.idc.fr](http://www.idc.fr)

### Copyright

**Publication externe des données et des informations d'IDC – toute information d'IDC destinée à être utilisée dans le cadre de publicités, de communiqués de presse ou de supports promotionnels doit préalablement faire l'objet du consentement écrit du directeur du bureau local d'IDC. Un projet de document proposé doit accompagner une telle demande. IDC se réserve le droit de refuser toute utilisation externe, quelle qu'en soit la raison. Copyright 2022 IDC. Toute reproduction sans autorisation écrite est strictement interdite. Tous droits réservés.**