Insee Première

N° 1643

Mars 2017

Cloud computing, *big data* : de nouvelles opportunités pour les sociétés

En 2016, parmi les sociétés de 10 personnes ou plus implantées en France, 17 % achètent des services de *cloud computing*, contre 12 % en 2014. En France, le recours au *cloud* demeure moins répandu qu'au niveau européen (21 %). Les plus grandes sociétés, celles de 250 personnes ou plus, sont les plus nombreuses à l'utiliser et leur recours s'est fortement accru entre 2014 et 2016. 48 % d'entre elles sont utilisatrices en 2016, contre 36 % en 2014. Le stockage de fichiers et la messagerie restent les services les plus fréquemment consommés mais les services de niveau avancé sont davantage achetés qu'en 2014.

Le recours au *cloud* privé, *via* des serveurs dédiés, augmente au détriment des serveurs partagés.

Par ailleurs, en 2015, 11 % des sociétés ont traité des données massives (*big data*). Près de deux tiers d'entre elles exploitent des données massives géolocalisées. Les sources mobilisées et la finalité des traitements dépendent beaucoup du secteur d'activité. Les trois quarts des sociétés qui exploitent des *big data* font réaliser certains traitements par leurs employés.

Enfin, la moitié des sociétés ont émis des factures électroniques. Seule une minorité de ces factures est dans un format permettant son traitement automatique.

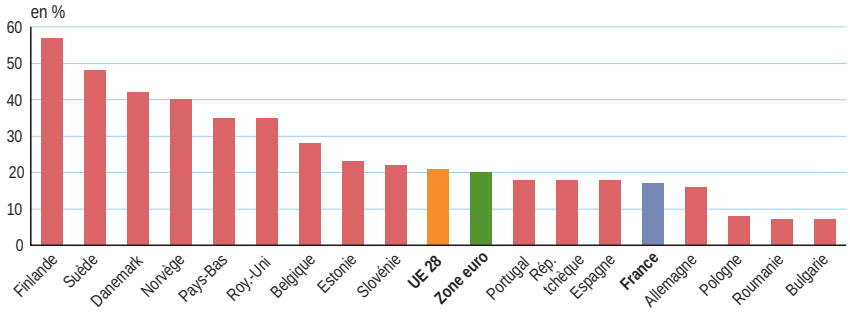
Thomas Vacher, Nadège Pradines, division Enquêtes thématiques et études transversales, Insee

En France, en 2016, 17 % des sociétés de 10 personnes ou plus *(sources)* achètent des services de *cloud computing* *(définitions)*, contre 12 % en 2014. Cette technologie permet d'exploiter *via* Internet la capacité de stockage ou la puissance de calcul de serveurs informatiques distants. En Allemagne, 16 % des sociétés utilisent des services de *cloud* payants. Ces services sont davantage utilisés dans les pays nordiques (plus de 40 % en Finlande, Suède, Danemark et Norvège ; *figure 1*).

L'usage du *cloud* payant est plus courant dans les grandes sociétés. Ainsi, en 2016, 48 % des sociétés de 250 personnes ou plus achètent ce type de services contre 21 % pour les sociétés de 20 à 249 personnes et 13 % pour celles de 10 à 19 personnes *(figure 2)*. Depuis 2014, le recours au *cloud* a surtout progressé dans les grandes sociétés (250 personnes ou plus). Dans les sociétés de plus petite taille, son usage augmente également, mais sans réduire leur écart avec les plus grandes.

Le secteur de l'information-communication reste le plus consommateur de *cloud* payant : en 2016, 51 % des sociétés de ce secteur achètent ces services. Viennent ensuite les sociétés des activités spécialisées, scientifiques et techniques (33 %). À l'opposé, seules 10 % des sociétés de l'hébergement-restauration les utilisent.

1 Sociétés utilisatrices de services payants de *cloud computing* en Europe en 2016



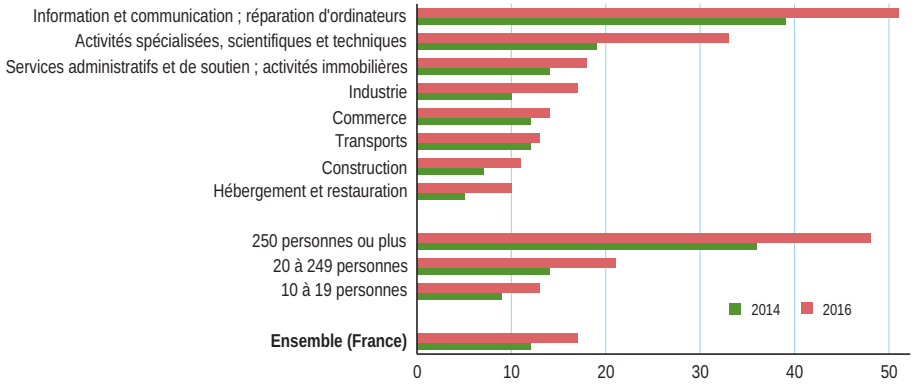
Lecture : en 2016, en Finlande, 57 % des sociétés consommaient des services de *cloud computing* payants.
Champ : sociétés de 10 personnes ou plus des secteurs principalement marchands hors secteurs agricole, financier et d'assurance, implantées dans l'Union européenne.
Source : Insee - Eurostat, enquête TIC 2016.

Les usages se sont diversifiés

En 2016, le stockage de fichiers devient le type de services le plus utilisé par les sociétés consommatrices de *cloud* payant (69 % d'entre elles, après 61 % en 2014 ; *figure 3*). Cet usage devance désormais celui du courriel (61 % après 62 % en 2014). La part de tous les autres services a cependant progressé.

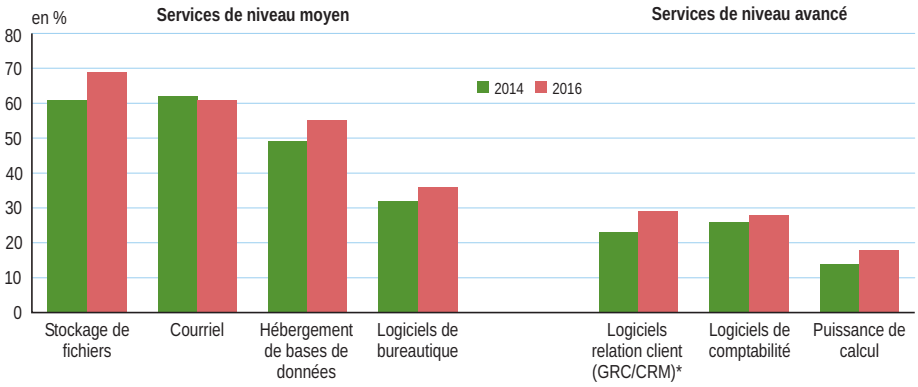
Insee

2 Sociétés utilisatrices de services payants de *cloud computing* en France en 2014 et 2016



Lecture : en 2014, 39 % des sociétés de l'information-communication et de la réparation d'ordinateurs consommaient des services de *cloud computing* payants. En 2016, 51 % de ces sociétés en consomment.
Champ : sociétés de 10 personnes ou plus des secteurs principalement marchands hors secteurs agricole, financier et d'assurance, implantées en France.
Source : Insee, enquêtes TIC 2014 et 2016.

3 Usages du *cloud computing* payant en France entre 2014 et 2016



* GRC/CRM : application pour la gestion de la relation client. Progiciel permettant de gérer l'ensemble des relations client dans un même processus.
Lecture : en 2014, 61 % des sociétés payant pour du *cloud computing* achetaient des services de stockage de fichiers. En 2016, 69 % en achètent (voir données complémentaires en ligne pour les résultats par secteur).
Champ : sociétés payant pour des services de *cloud computing*, de 10 personnes ou plus, des secteurs principalement marchands hors secteurs agricole, financier et d'assurance, implantées en France.
Source : Insee, enquêtes TIC 2014 et 2016.

La consommation de services de courriel et de stockage, l'hébergement de bases de données ou l'utilisation de logiciels de bureautique sont des ressources auxquelles les sociétés peuvent accéder sans recourir au *cloud* payant. Ces services impliquent *a priori* une dépendance limitée de la société vis-à-vis du *cloud*. D'autres services de *cloud*, de niveau avancé, sont moins utilisés mais se développent également : l'achat de logiciels de comptabilité, de logiciels de relation client et le recours à de la puissance de calcul. Ces services moins facilement substituables induisent une dépendance plus forte des sociétés clientes au *cloud* payant.

Près de la moitié des sociétés consommatrices de *cloud* utilisent uniquement des services impliquant une dépendance limitée vis-à-vis du *cloud*, soit 7 points de moins qu'en 2014. Cette baisse s'explique par un recours accru à des services avancés. Ainsi, une société consommatrice de *cloud* sur deux achète au moins un service de niveau avancé, soit 6 points de plus qu'en 2014.

Le *cloud* privé progresse

Les services de *cloud* payant peuvent être fournis soit par des serveurs informatiques partagés (*cloud* public) soit par des serveurs exclusivement dédiés à un client (*cloud* privé). En 2016, 58 % des sociétés de 10 personnes ou plus qui utilisent le *cloud* le font *via* des serveurs dédiés, 59 % *via* des serveurs partagés et 17 % ont recours aux deux types de *cloud*. Depuis 2014, la part de sociétés utilisant conjointement les deux types de serveurs n'a pas évolué et le recours au *cloud* privé augmente au détriment des serveurs partagés.

Dans les deux secteurs utilisant le moins de serveurs privés en 2014, le *cloud* partagé est désormais minoritaire : ainsi, dans l'hébergement-restauration, en 2016, 55 % des sociétés font appel exclusivement à des serveurs privés, contre seulement 22 % en 2014. Dans les transports, cette proportion est de 42 % en 2016 contre 27 % en 2014. En 2016, seule la construction utilise moins de serveurs privés qu'en 2014.

Avant l'émergence des services de *cloud computing*, les traitements de données informatiques étaient principalement réalisés au sein d'entrepôts de données propres à l'entreprise. Par les capacités de stockage et de puissance de calcul accrues qu'il offre, le *cloud* peut apporter une solution aux sociétés souhaitant traiter des données de plus en plus abondantes et volumineuses produites par les systèmes d'information. Ainsi, à taille et secteur d'activité donnés, la consommation de services de *cloud* payant est deux fois plus fréquente chez les sociétés qui font des analyses de données massives (*big data* ; *définitions)* que chez celles qui n'en font pas.

4 Sociétés utilisatrices de *big data* et sources utilisées, en 2015

Secteur (niveau A17 de la NA)	Sociétés utilisatrices de *big data*	Sources utilisées par les sociétés qui procèdent à l'analyse de *big data*			
		Géolocalisation	Médias sociaux	Objets connectés/capteurs	Autres
Transports	25	92	7	20	4
Information et communication ; réparation d'ordinateurs	23	33	64	38	26
Services administratifs et de soutien ; activités immobilières	13	75	34	14	8
Construction	11	89	13	14	4
Activités spécialisées, scientifiques et techniques	11	43	43	36	18
Commerce	9	50	34	35	8
Hébergement et restauration	9	49	76	25	16
Industrie	8	51	25	46	10
Ensemble (France)	**11**	**62**	**32**	**29**	**10**

Lecture : 25 % des sociétés des transports exploitent des *big data* ; parmi elles, 92 % exploitent des données de sources géolocalisées.
Champ : sociétés d'au moins 10 personnes des secteurs principalement marchands hors secteurs agricole, financier et d'assurance, implantées en France.
Source : Insee, enquête TIC 2016.

11 % des sociétés ont analysé des *big data*

En 2015, 11 % des sociétés de 10 personnes ou plus ont procédé à une analyse de *big data* *(figure 4)*. Ces données présentent généralement d'importants volumes et une variété de formats complexes. On utilise en général le terme de *big data* pour les données dont le volume, l'actualisation constante et la structure rendent difficile l'exploitation par des outils classiques de gestion ou d'analyse de bases de données, et qui requièrent donc l'utilisation de technologies spécifiques.
Le traitement de données massives est davantage pratiqué par les grandes sociétés. Parmi celles de 250 personnes ou plus, 24 % ont exploité des *big data* contre 14 % pour celles de 20 à 249 personnes et 9 % de celles de 10 à 19 personnes. Les sociétés des transports (25 %) et de l'information-communication (23 %) utilisent le plus des *big data*. Les sociétés de l'industrie y ont le moins souvent recours (8 %).

Les données massives géolocalisées sont les plus utilisées

Les *big data* peuvent être issues des propres données de la société provenant d'objets intelligents ou connectés, de capteurs, de données de géolocalisation (adresse IP, antenne relais de téléphonie mobile, etc.) ou encore de données générées par les médias sociaux. Les données de géolocalisation constituent la première source de *big data* : 62 % des sociétés exploitant des *big data* traitent ce type de données (c'est-à-dire 7 % de l'ensemble des sociétés de 10 personnes ou plus). En outre, 32 % des utilisatrices de *big data* exploitent des données issues des médias sociaux (4 % de l'ensemble) et 29 % des informations relevées par des objets intelligents, connectés ou des capteurs (3 % de l'ensemble).

Parmi les sociétés analysant des *big data*, les données géolocalisées sont très largement mobilisées dans les transports (92 %) et la construction (89 %). L'utilisation des sources issues des médias sociaux concerne 76 % des sociétés utilisatrices dans l'hébergement-restauration et 64 % dans l'information-communication. Enfin, l'analyse de données à partir d'objets connectés est surtout utilisée par les sociétés de l'industrie (46 %).

Selon les sources de données utilisées, la finalité est différente

Les finalités de traitement sont multiples et réparties de façon équilibrée. Ainsi, pour 38 % des sociétés réalisant des analyses de *big data*, le but est d'optimiser des processus internes de production de biens ou de fourniture de services. Elles exploitent alors majoritairement des données de géolocalisation.
Par ailleurs, pour 38 % des sociétés utilisatrices, le but est de développer ou d'améliorer des biens ou services. Elles recourent autant aux données de géolocalisation qu'à celles d'objets connectés et des médias sociaux.
Enfin, 35 % des sociétés cherchent à améliorer le marketing ou la gestion des ventes. Pour cela, les données des médias sociaux sont majoritairement mobilisées, suivies par les données géolocalisées.

Le traitement des *big data* est majoritairement réalisé en interne

Dans les sociétés analysant des *big data*, 74 % font réaliser les traitements par leurs employés *(figure 5)* et 42 % par des prestataires externes. 16 % d'entre elles recourent aux deux modes. La nature des données à traiter n'a pas d'incidence sur le mode de traitement.
Dans l'information-communication et les activités spécialisées, scientifiques et techniques, les données sont davantage traitées en interne, car les employés sont probablement mieux formés pour cela que dans d'autres secteurs. Ainsi, parmi celles qui utilisent le *big data*, 90 % des sociétés de l'information-communication réalisent ces traitements en interne, 84 % dans les activités spécialisées, scientifiques et techniques. Même si le traitement en interne est le plus courant, le commerce et l'hébergement-restauration font plus souvent appel à des prestataires externes. Dans l'hébergement-restauration, 41 % des sociétés qui exploitent des *big data* le font uniquement par l'intermédiaire d'un prestataire externe, contre 10 % des sociétés de l'information-communication.
Tous secteurs confondus, 83 % des sociétés de 250 personnes ou plus font réaliser les traitements exclusivement par leurs employés, contre 73 % des sociétés de 10 à 249 personnes.

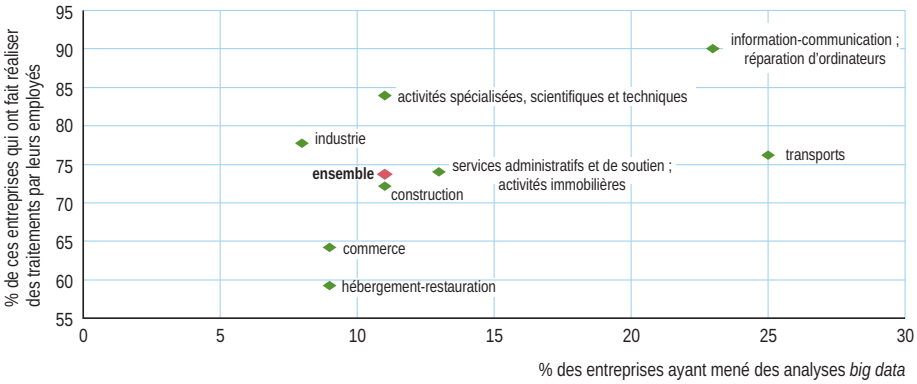
La méconnaissance des *big data*, principal facteur de non utilisation

Différents facteurs, parfois cumulés, peuvent empêcher ou limiter l'utilisation des *big data* : 53 % des sociétés évoquent une méconnaissance du sujet ou n'en voient pas l'intérêt. 23 % font état d'un manque de compétences et 21 % trouvent le coût de traitement trop élevé. En outre, 10 % des sociétés mettent en avant une qualité des résultats insuffisante et 6 % jugent contraignante la législation sur la protection des données.
Les grandes sociétés (250 personnes ou plus) se déclarent davantage limitées ou empêchées par le cadre législatif (10 % d'entre elles). Elles sont au contraire moins limitées par la méconnaissance du sujet (44 % d'entre elles).

La moitié des sociétés ont émis des factures électroniques

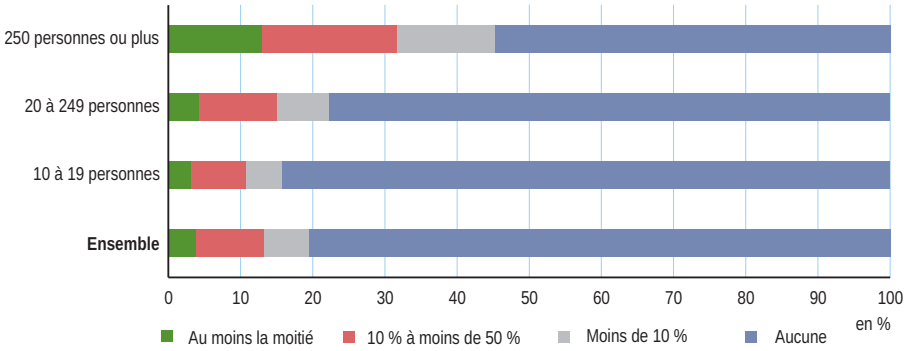
Au cours de l'année 2015, les trois quarts des sociétés ont émis des factures à d'autres sociétés ou administrations publiques. 70 % de ces sociétés ont expédié des factures sous un format électronique, soit la moitié de l'ensemble des sociétés.
Pour 19 % des sociétés émettrices de factures, les factures envoyées peuvent être traitées automatiquement. Même parmi les sociétés dont les factures émises peuvent être traitées automatiquement, ce type d'envoi reste minoritaire : 32 % d'entre elles en ont envoyé au moins un dixième dans un format pouvant être traité automatiquement et seules 20 % en envoient plus de la moitié dans ce format.
Ces envois sont plus répandus dans les sociétés de 250 personnes ou plus émettant des factures : 32 % d'entre elles en ont émis plus d'un dixième dans un format pouvant être traité automatiquement, contre 15 % des sociétés de 20 à 249 personnes et 11 % des sociétés de 10 à 19 personnes *(figure 6)*. ■

5 Traitement des *big data* par les employés de la société en 2015



Lecture : 23 % des entreprises de l'information-communication et de la réparation d'ordinateurs ont mené des analyses *big data*. 90 % de ces entreprises ont eu recours à leurs employés pour certains traitements.
Champ : sociétés de 10 personnes ou plus des secteurs principalement marchands hors secteurs agricole, financier et d'assurance, implantées en France.
Source : Insee, enquête TIC 2016.

6 Factures émises permettant un traitement automatique en 2016 selon la taille des sociétés



Lecture : 19 % des sociétés ont émis des factures électroniques pouvant être traitées automatiquement. Pour 4 % des sociétés, les factures de ce format représentent au moins la moitié des factures émises.
Champ : sociétés d'au moins 10 personnes des secteurs principalement marchands hors secteurs agricole, financier et d'assurance, implantées en France et ayant émis des factures pour d'autres sociétés ou pour l'administration publique.
Source : Insee, enquête TIC 2016.

Sources

L'**enquête sur les technologies de l'information et de la communication et le commerce électronique (TIC)** de 2016 a été réalisée début 2016 auprès d'un échantillon de 12 700 unités légales, sociétés ou entreprises individuelles, actives, occupant 10 personnes ou plus (salariés ou non-salariés) et implantées en France. Le terme « sociétés » est utilisé ici pour les désigner de manière générique. Cette étude s'appuie en effet sur la définition juridique de l'entreprise (unité légale), et non sur la définition économique instaurée par la loi de modernisation de l'économie (LME) et son décret d'application n° 2008-1354 du 18 décembre 2008.
L'enquête porte sur les secteurs d'activité suivants : industrie, construction, commerce et réparation d'automobiles et de motocycles, transports et entreposage, hébergement-restauration, information-communication et réparation d'ordinateurs, activités spécialisées, scientifiques et techniques, activités de services administratifs et de soutien, et activités immobilières (sections C à J, L, M hors 75, N et groupe 95.1 de la NAF rév. 2). L'échantillon est représentatif d'environ 191 000 sociétés.
Parmi les sociétés actives de l'échantillon, 82 % ont répondu à l'enquête. En matière de précision, la proportion de sociétés ayant recours à des services de *cloud*, par exemple, est estimée à 17 % avec un intervalle de confiance à 95 % de 1 point.
L'enquête vise à mieux connaître l'informatisation et la diffusion des TIC dans les entreprises. Les questions sur les taux d'équipement en TIC portent en général sur la situation au moment de l'enquête, c'est-à-dire au cours du premier trimestre 2016. Celles portant sur certaines pratiques, notamment les *big data*, se réfèrent à l'année précédant l'enquête (soit 2015 pour l'enquête 2016). En 2016, le questionnaire comporte deux modules spécifiques de questions sur le *cloud computing*, dont certaines avaient été posées en 2014, et l'usage des données massives (*big data*). Des enquêtes analogues ont été menées dans tous les pays européens en application du règlement communautaire n° 1006/2009 sur la société de l'information. Leurs résultats sont disponibles sur le site d'Eurostat et s'interprètent à la lumière des rapports qualité également fournis (http://ec.europa.eu/eurostat/cache/metadata/en/isoc_e_esms.htm).

Définitions

Le ***cloud computing***, appelé aussi **informatique en nuage** ou ***cloud***, désigne des services informatiques utilisés sur Internet, délivrés par les serveurs d'un prestataire. Il permet à une société d'accéder à des ressources informatiques sur Internet (stockage, puissance de calcul, logiciels) plutôt que *via* sa propre infrastructure informatique.
Les services de *cloud* pris en compte dans cette étude sont payants, mobilisables par l'utilisateur à la demande et en libre service, c'est-à-dire sans besoin d'intervention du prestataire une fois installés, et modulables selon les besoins de la société.
Les ***big data***, aussi appelées **données massives** ou **mégadonnées**, sont générées par les activités exécutées électroniquement et entre machines (par exemple, données issues des médias sociaux, des processus de production, etc.). Les *big data* présentent généralement les caractéristiques suivantes, dites des « 3 V » :
- un volume important issu d'une grande quantité de données générées au cours du temps ;
- une variété concernant les formats de données complexes, structurées ou non (par exemple, textes, vidéos, documents, données de capteurs, flux de clics, coordonnées, etc.) ;
- la vélocité du fait de la grande vitesse à laquelle les données sont générées, deviennent disponibles et évoluent au cours du temps.

Bibliographie

- « Les technologies de l'information et de la communication et le commerce électronique 2016 », *Insee Résultats,* à paraître.
- Demoly E., Vacher T., « Sécurité numérique et médias sociaux dans les entreprises en 2015 », *Insee Première* n° 1594, mai 2016.
- Vacher T., Demoly E., « La timide émergence du *cloud computing* dans les sociétés en 2014 », *Insee Première* n° 1545, avril 2015.

Direction Générale :
18, bd Adolphe-Pinard
75675 PARIS CEDEX 14
Directeur de la publication :
Jean-Luc Tavernier
Rédacteur en chef :
E. Nauze-Fichet
Rédacteurs :
J.-B. Champion, C. Collin, C. Lesdos-Cauhapé, V. Quénechdu, H. Valdelièvre
Maquette : B. Rols
Impression : Jouve
Code Sage IP171643

ISSN 0997 - 3192

- ***Insee Première*** figure dès sa parution sur le site internet de l'Insee :
https://www.insee.fr/fr/statistiques?collection=116
- Pour recevoir par courriel les avis de parution (50 numéros par an) :
https://www.insee.fr/fr/information/1405555

Pour vous abonner à ***Insee Première*** et le recevoir par courrier :
https://www.insee.fr/fr/information/2537715

Insee
Mesurer pour comprendre