
Dossier



Avertissement

Sur les sites internet de l'Insee (www.insee.fr) pour les données françaises et d'Eurostat (<https://ec.europa.eu/eurostat>) pour les données internationales, les chiffres essentiels sont actualisés régulièrement.

Les données chiffrées sont parfois arrondies, en général au plus près de leurs valeurs réelles. Le résultat arrondi d'une combinaison de chiffres (qui fait intervenir leurs valeurs réelles) peut être légèrement différent de celui que donnerait la combinaison de leurs valeurs arrondies.

Les comparaisons internationales contenues dans cet ouvrage s'appuient sur des données harmonisées publiées par Eurostat, qui peuvent différer des données nationales diffusées par les instituts nationaux de statistique.

Cloud computing et big data : la dématérialisation au service des sociétés européennes

Nadège Pradines*

Depuis les années 2000, le *cloud computing*, ou *cloud* (informatique en nuage), permet aux sociétés de s'affranchir de tout ou partie de leurs investissements en infrastructures informatiques pour utiliser, à la place, des services en ligne : logiciels, stockage, puissance de calcul, etc. Le recours au *cloud* payant est adopté par une société de 10 personnes ou plus sur quatre dans l'Union européenne (UE) – une sur cinq en France en 2018 –, cet écart étant principalement dû aux sociétés de moins de 250 personnes.

Les services achetés sur le *cloud* diffèrent selon la taille des sociétés : les petites sociétés achètent davantage des logiciels de comptabilité et les grandes, des solutions de gestion de la relation client. L'achat de puissance de calcul sur le *cloud*, aisément modulable à la hausse ou à la baisse, facilite le recours à une nouvelle pratique : l'analyse de données massives (*big data*). 16 % des sociétés de 10 personnes ou plus en analysent en France, contre 12 % en moyenne dans l'UE. En France, les sociétés réalisant des analyses de données massives utilisent davantage des données issues de la géolocalisation d'appareils portables qu'en moyenne dans l'UE.

L'entrée dans l'ère numérique des entreprises a créé un besoin fort d'infrastructures informatiques. La maintenance des infrastructures informatiques est la tâche informatique la plus réalisée : en 2018, en France, elle concerne 85 % des sociétés¹ de 10 personnes ou plus. N'étant pas, la plupart du temps, une tâche du cœur de métier, elle est aussi la tâche informatique la plus sous-traitée : 62 % des sociétés ont principalement recours à un prestataire quand seulement 23 % font réaliser cette maintenance par leurs propres employés.

Dans l'Union européenne, une société de 10 personnes ou plus sur quatre a recours au *cloud* payant

La gestion externalisée des infrastructures informatiques est la caractéristique principale du *cloud computing* ou *cloud* (informatique en nuage). Il consiste à fournir *via* Internet des services informatiques aux entreprises, leur permettant ainsi de réaliser des économies et de transformer des investissements physiques (serveurs, réseau, etc.) en contrats de location de services en ligne. Le *cloud* offre différents niveaux de service : il propose l'accès à des logiciels (*software as a service*, SaaS), de la puissance de calcul, une capacité de stockage, etc. Délivré par les serveurs informatiques du fournisseur, il est facilement modulable à la hausse ou à la baisse, pour s'adapter en temps réel aux besoins de l'entreprise. L'enquête communautaire TIC-entreprises (*encadré 1*) mesure l'usage payant de ces services, sachant qu'il existe des offres gratuites, non mesurées.

* Nadège Pradines (Insee).

1. Dans les pays de l'UE, y compris la France, l'enquête communautaire TIC-entreprises rassemble à ce jour une information sur des unités légales de 10 personnes ou plus sans considération sur leur appartenance à un groupe. Nous les désignons ici par le terme « sociétés ».

Enquête communautaire TIC-entreprises

L'enquête sur les technologies de l'information et de la communication et le commerce électronique auprès des entreprises (TIC-entreprises / *ICT survey in enterprises*) est menée tous les ans dans les pays de l'Union européenne auprès des sociétés des secteurs principalement marchands, hors secteurs financier, d'assurance et agricole. Elle s'adresse aux sociétés de 10 personnes ou plus. Certains thèmes reviennent selon une périodicité variable. Le *cloud computing* payant a été abordé pour la première fois en 2014, puis en 2016 et 2018. L'analyse de données massives a été abordée pour la première fois en 2016 (sur les analyses réalisées en 2015), puis une nouvelle fois dans l'enquête 2018 (sur les analyses réalisées en 2017).

Quand un thème est optionnel au niveau européen, certains pays choisissent de ne pas l'inclure dans leur enquête : cela conduit à l'absence de données sur l'analyse de données massives en 2016 pour l'Autriche, la Lettonie et l'Irlande, et en 2018 pour le Royaume-Uni. Enfin, au cas par cas, certains pays jugent que la qualité d'une variable n'est pas suffisante pour la diffuser, ce qui peut causer l'absence ponctuelle d'informations sur un pays. Les valeurs données pour la « moyenne dans l'Union européenne » incluent le plus souvent les valeurs masquées à cause de leur qualité, mais excluent en revanche les pays n'ayant pas fourni de données.

Dans l'Union européenne² (UE), 26 % des sociétés payent pour des services de *cloud computing* en 2018. Les pays où les sociétés sont les plus nombreuses à en consommer sont les trois pays nordiques de l'UE : Finlande (65 %), Suède (57 %) et Danemark (56 %) (*figure 1*). Leur avance numérique n'est pas propre à l'utilisation du *cloud* et concerne également les usages numériques des particuliers. En effet, dans ces pays, les technologies numériques se sont déployées tôt et vite, encouragées par une numérisation forte des services publics et un contexte économique spécifique, notamment en Finlande : l'importance de Nokia puis son échec sur le marché des *smartphones* et l'essaiage de ses anciens employés dans le tissu économique finlandais ont contribué à la forte numérisation des entreprises. De plus, les pays du Nord ont un avantage géographique en matière de *cloud* : leur température moyenne est favorable au refroidissement des grands centres de serveurs de données (*data centers*) dont se dotent les fournisseurs de *cloud*.

En France, une société sur cinq utilise le *cloud* : un écart imputable aux sociétés de moins de 250 personnes

En 2018, l'usage du *cloud* est encore peu développé en France : seuls 19 % des sociétés de 10 personnes ou plus payent pour des services de *cloud computing*. Les pays où l'usage de *cloud computing* est moindre sont des pays d'Europe de l'Est (Hongrie, Lettonie, Pologne, Roumanie, Bulgarie) et la Grèce (*encadré 2*).

Si l'écart entre la France et l'UE était déjà de 7 points en 2014 (12 % contre 19 % dans l'UE), comme en 2018, il semblait s'être réduit en 2016, avec 17 % des sociétés payant pour du *cloud* contre 21 % dans l'UE. Dans tous les cas, le recours au *cloud computing* a augmenté pour tous les pays entre 2014 et 2018³. Cette progression a été la plus forte aux

2. Dans ce dossier, il s'agit de l'Union européenne à 28 pays dans ses frontières de 2013 à 2018.

3. Pour l'Italie, la valeur communiquée dans les données pour 2014 n'est pas comparable avec les millésimes suivants (anormalement haute). Après consultation de l'*Istituto nazionale di statistica*, l'évolution 2015-2018 est commentée ici.

Pays-Bas : + 21 points en quatre ans. En 2018, 48 % des sociétés néerlandaises payent des services de *cloud computing*.

Alors que 23 % des sociétés de 10 à 49 personnes de l'UE payent pour des services de *cloud computing* (16 % en France), elles sont 56 % parmi les grandes sociétés (de 250 personnes ou plus). Malgré son retard global en matière de recours au *cloud computing*, la France est au-dessus de la moyenne européenne quand sont considérées les seules grandes sociétés : 60 % d'entre elles payent pour des services de *cloud*. Ce fort recours des grandes sociétés en France explique peut-être l'intérêt des principaux fournisseurs de *cloud* pour le territoire français : si Google n'a pas encore de *data center* en France, Microsoft a ouvert les siens en décembre 2017 et Amazon Web Services (AWS) en janvier 2018 (encadré 3).

Encadré 2

Les difficultés de comparaisons entre les pays

Le règlement européen auquel répondent les enquêtes TIC-entreprises décrit précisément les indicateurs à fournir mais ne pose aucune obligation sur le mode de collecte. Ainsi, certains pays fournissent les chiffres d'affaires renseignés par les enquêtés, d'autres fournissent les chiffres d'affaires obtenus *via* des sources fiscales. Dans la mise en œuvre de l'enquête communautaire, des écarts dans les pratiques et méthodes peuvent conduire à des différences entre pays :

- l'adaptation nationale du questionnaire : traduction, changement de l'ordre des modules, ajout ou retrait de questions, altération des filtres (les données 2014 sur le *cloud computing* en Italie ne sont pas comparables à celles collectées depuis à cause de différences dans la formulation du questionnaire national) ;

- le dispositif légal national : selon les pays, l'enquête peut être obligatoire ou sur la base du volontariat (enquête non obligatoire en Allemagne, en Irlande, en Autriche, en Belgique et aux Pays-Bas).

Une enquête obligatoire permet, en général, un meilleur taux de réponse. Elle force à la réponse des entreprises qui ne se sentent pas concernées par les sujets de l'enquête (parce qu'elles font un usage moindre des TIC, par exemple). Ce faisant, la population décrite par les résultats de l'enquête est plus représentative de l'ensemble des entreprises.

Par exemple, en Allemagne, le taux de réponse est de 31 % (contre 81 % ou plus dans la moitié des pays de l'Union européenne). Il est probable que les résultats de cette enquête souffrent d'un biais de participation : ils peuvent surestimer

le développement des TIC dans les entreprises parce que les sociétés répondant volontairement sont probablement plus concernées par les TIC.

- L'échantillonnage : dans sa stratification, la France ou la République tchèque incluent le secteur, l'effectif et le chiffre d'affaires, ce dernier critère augmentant la représentativité économique des entreprises. Au contraire, des pays comme l'Allemagne, la Grèce, l'Espagne ou la Belgique utilisent plutôt le secteur, l'effectif et la région, ce qui permet la production de résultats régionaux mais diminue la représentativité de l'activité économique. D'autres pays croisent seulement secteur et effectif (Estonie, Danemark, Finlande, Bulgarie, etc.).

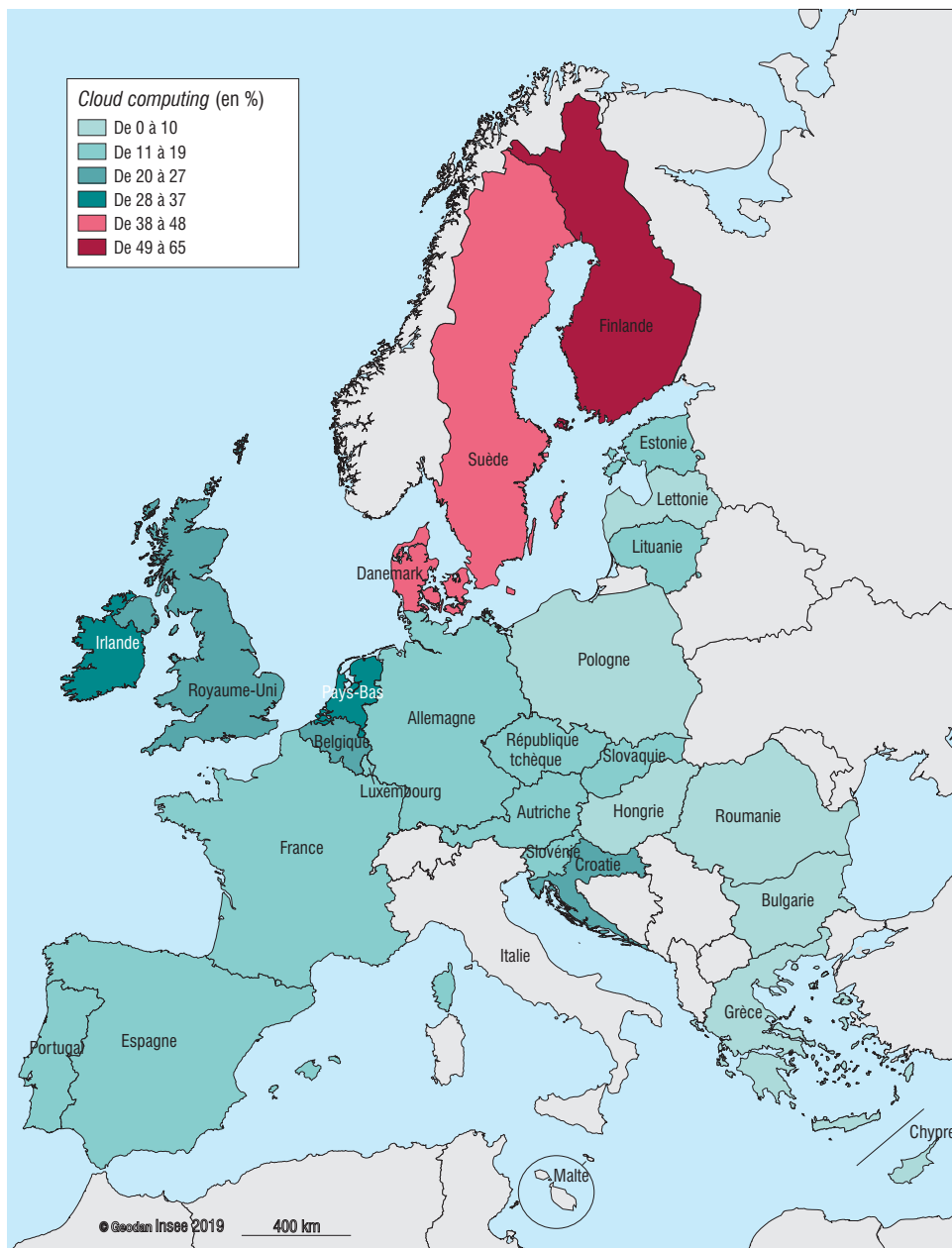
- La collecte : dans presque tous les pays, il était possible de répondre par Internet en 2018. Il s'agit pour certains de questionnaires en ligne ; pour d'autres, de télécharger le questionnaire PDF, le remplir et le retourner par courriel. Cela peut entraîner des écarts liés aux modes de collecte.

- Les traitements post-collecte : certains pays ne contrôlent pas la cohérence des données à l'issue de la collecte ou ne prennent pas en charge la correction de la non-réponse partielle. En 2018, il s'agit notamment du Luxembourg, de la Finlande, du Danemark, de la Suède, de la Lettonie et de l'Estonie. Pour ces pays, les données produites sont donc possiblement sous-estimées.

Pour toutes ces raisons, les comparaisons commentées dans ce dossier doivent être prises avec précaution et les faibles écarts ne doivent pas être surinterprétés.

1. Évolution du recours au *cloud computing* payant entre 2014 et 2018 dans l'UE

a. Recours au *cloud computing* payant en 2014 dans l'UE



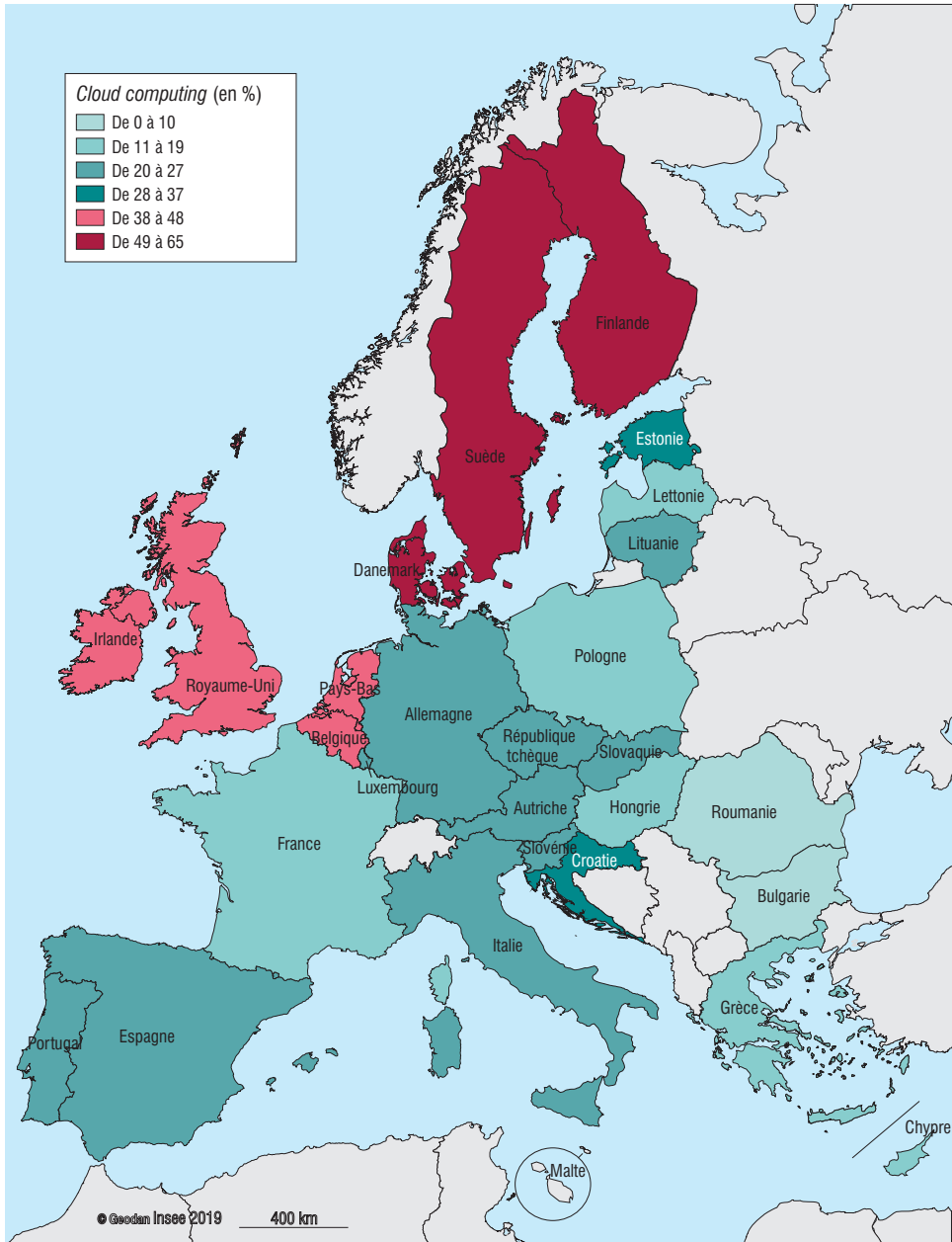
Note : en 2014, les données sur le *cloud computing* pour l'Italie ne sont pas disponibles.

Lecture : en 2014, entre 11 % et 19 % des sociétés au Portugal achètent des services de *cloud computing*.

Champ : sociétés de 10 personnes ou plus des secteurs principalement marchands, hors secteurs agricole, financier et d'assurance.

Sources : Eurostat ; Insee, enquêtes TIC-entreprises 2014.

b. Recours au *cloud computing* payant en 2018 dans l'UE



Lecture : en 2018, entre 20 % et 27 % des sociétés au Portugal achètent des services de *cloud computing*.
 Champ : sociétés de 10 personnes ou plus des secteurs principalement marchands, hors secteurs agricole, financier et d'assurance.
 Sources : Eurostat ; Insee, enquêtes TIC-entreprises 2018.

Enjeux et limites du *cloud computing*

Les trois principaux fournisseurs de services de *cloud computing* sont Amazon Web Services (AWS), Google et Microsoft. À côté de leur offre payante, ces grands fournisseurs proposent également du *cloud computing* gratuit (par exemple Google Mail, Docs, Drive, etc.), qui est peut-être privilégié par les sociétés de taille moindre et échappe à la mesure par l'enquête TIC-entreprises.

En dépit de l'absence virtuelle de frontières dans le domaine du numérique, la localisation des centres de données peut influencer le recours au *cloud computing*, de même que la nationalité du fournisseur de services. En effet, le pays où sont stockées les données peut faire l'objet de certaines réglementations, comme le *Patriot Act*, qui permet aux autorités américaines de prendre connaissance de toutes données hébergées par une société sur le sol américain. Quant au *Cloud Act*, entré en vigueur en 2018, il permet aux autorités américaines d'avoir accès à toutes les données hébergées par une

société de droit américain, même hors du territoire des États-Unis et en dépit des lois européennes sur la protection des données individuelles. Face à ces risques, des offres de *cloud* dit « souverain » existent, qui garantissent que les centres de données sont implantés dans le même pays que les sociétés qui payent pour ce *cloud* et relèvent du droit national. Une note d'information publiée par le ministère de l'Intérieur français précise que « l'utilisation d'un *cloud* non souverain [...] est illégale pour toute institution produisant des archives publiques, dont les collectivités territoriales, leurs groupements et leurs établissements publics ».

Enfin, confier ses données à un prestataire comporte des risques en matière de sécurité et d'intégrité des informations hébergées. En France, l'Agence nationale de sécurité des systèmes d'information (ANSSI) a établi un référentiel, SecNumCloud, qui permet de certifier des prestataires *cloud* de confiance.

Plus de recours à l'hébergement de bases de données et de fichiers en France que dans l'UE

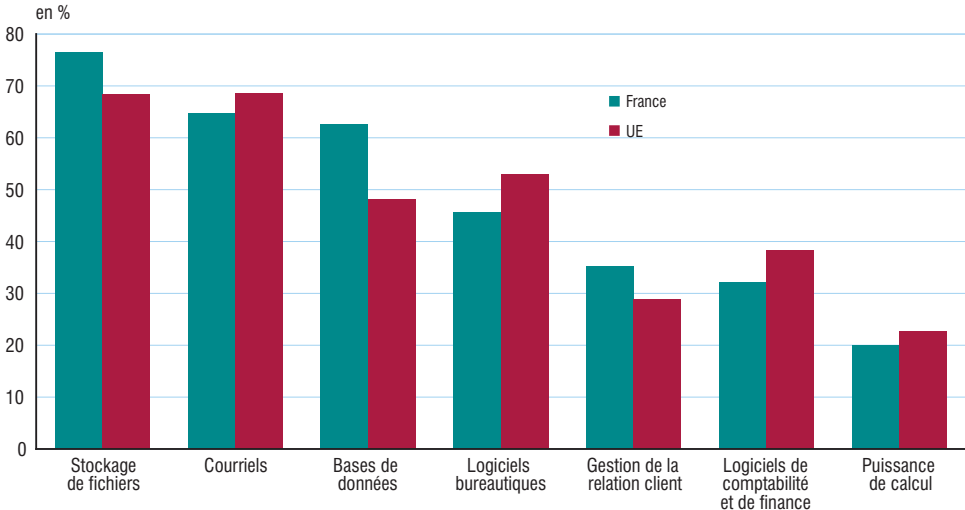
L'usage de *cloud computing* payant recouvre des types de services différents. L'enquête TIC-entreprises distingue les services suivants : le courriel sur le *cloud*, les logiciels de bureautique sur le *cloud* et le stockage de fichiers de l'entreprise, considérés comme des services de faible niveau technologique ; l'hébergement de bases de données est considéré comme un service de niveau intermédiaire pouvant nécessiter des compétences techniques ; enfin, l'exécution sur le *cloud* de logiciels de finance ou de comptabilité, de logiciels de gestion de la relation client ou l'achat de puissance de calcul sur le *cloud* sont considérés comme des services de haut niveau, à la fois pour les compétences nécessaires à l'utilisation de ces services et le service rendu. Le *cloud* ne se limite pas à ces usages et peut fournir d'autres services, non précisés dans le questionnaire de l'enquête.

Les pratiques d'usage du *cloud* varient selon les pays. Ainsi, l'utilisation du *cloud computing* pour de l'hébergement de bases de données, à l'exclusion de tout autre service précisé, est particulièrement forte pour les sociétés qui payent du *cloud* en Espagne (22 %), en France (23 %) et en Bulgarie (26 %).

En France, 77 % des sociétés payant des services de *cloud* souscrivent à un service de stockage de fichiers et 63 % à de l'hébergement de bases de données, contre respectivement 68 % et 48 % dans l'UE. À l'inverse, l'utilisation de logiciels en *cloud* (SaaS) y est moindre : moins de courriel en *cloud* (65 % en France *versus* 69 % dans l'UE), de logiciels bureautiques en *cloud* (46 % contre 53 %), de logiciels de comptabilité et de puissance de calcul. Seuls les services de gestion de la relation client sont plus utilisés, proportionnellement, par les sociétés en France recourant au *cloud* (figure 2).

En Grèce et en Autriche, les sociétés payant pour du *cloud computing* souscrivent davantage à des services de faible niveau uniquement (respectivement 45 % et 46 %) qu'à des services de haut niveau (38 % et 40 %). Dans certains pays, non seulement les sociétés utilisent plus de *cloud* payant, mais elles souscrivent en général davantage à des services de haut niveau. Ainsi, aux Pays-Bas, 69 % des sociétés payant du *cloud* utilisent des services de haut niveau, contre 9 % pour des services de faible niveau uniquement (figure 3).

2. Types de services souscrits par les entreprises payant pour du cloud computing dans l'UE et en France en 2018

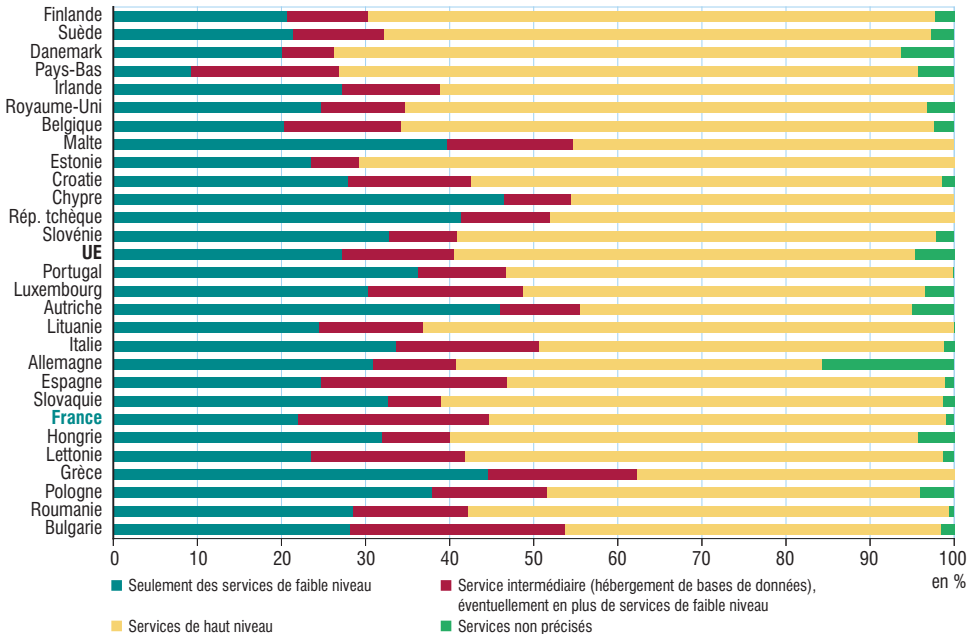


Lecture : en 2018, 77 % des sociétés de 10 personnes ou plus implantées en France et payant pour des services de *cloud computing* recourent au stockage de fichiers sur le *cloud* payant.

Champ : sociétés de 10 personnes ou plus des secteurs principalement marchands, hors secteurs agricole, financier et d'assurance, payant pour des services de *cloud computing* en 2018.

Sources : Eurostat ; Insee, enquête TIC-entreprises 2018.

3. Types de services de cloud computing payant utilisés dans les pays de l'UE en 2018



Note : les pays sont classés par ordre décroissant d'usage du *cloud computing* par les sociétés.

Lecture : en 2018, la Finlande est le pays de l'UE où la plus grande part de sociétés payent pour des services de *cloud computing*. Parmi elles, 21 % n'utilisent sur le *cloud* que des services de faible niveau et 67 % utilisent des services de haut niveau.

Champ : sociétés de 10 personnes ou plus des secteurs principalement marchands, hors secteurs agricole, financier et d'assurance, payant pour du *cloud computing* en 2018.

Sources : Eurostat ; Insee, enquête TIC-entreprises 2018.

Les petites sociétés utilisent plus de logiciels de comptabilité sur le cloud, les grandes sociétés plus de logiciels de gestion de la relation client

En France comme dans l'UE, quelle que soit leur taille, un peu plus de la moitié des sociétés payant pour du *cloud* utilisent des services de haut niveau. Cependant, ces services diffèrent : en France, 35 % des sociétés de 10 à 49 personnes payant du *cloud* le font pour accéder à des logiciels de comptabilité, contre seulement 20 % pour les sociétés de 250 personnes ou plus. À l'inverse, ces grandes sociétés payant pour le *cloud* souscrivent plus souvent à des logiciels de gestion de la relation client (43 %) et de la puissance de calcul (25 %) que les sociétés de 10 à 49 personnes payant du *cloud* (respectivement 34 % et 18 %). Les usages de services de niveau faible, quels qu'ils soient, dépendent moins de la taille de la société (figure 4).

Si le passage au *cloud* est *a priori* une question de choix pour les sociétés, il peut arriver que celles-ci soient contraintes par la migration d'une partie de l'offre de logiciels vers le *cloud*. Souscrire au *cloud* serait alors le seul moyen de pouvoir utiliser certains logiciels. C'est le cas de certains logiciels de caisse, dont l'offre de *cloud* s'étend au détriment de l'offre standard. Avec l'obligation, au 1^{er} janvier 2018, des logiciels de caisse certifiés, l'usage du *cloud computing* dans le commerce ou l'hébergement-restauration pourrait augmenter en France. Il est déjà passé de 5 % des sociétés de l'hébergement-restauration de 10 personnes ou plus en 2014 à 12 % en 2018.

4. Types de services souscrits par les entreprises payant pour du *cloud computing* en 2018 selon l'effectif

en %

	France			UE		
	250 personnes ou plus	50 à 249 personnes	10 à 49 personnes	250 personnes ou plus	50 à 249 personnes	10 à 49 personnes
Services de faible niveau						
Courriels	67	72	62	64	67	69
Logiciels bureautiques	56	53	42	58	55	52
Stockage de fichiers	77	74	77	70	68	68
Services intermédiaires						
Hébergement de bases de données	62	60	63	49	47	48
Services de haut niveau						
Logiciels de comptabilité et de finance	20	27	35	26	32	41
Gestion de la relation client	43	36	34	34	31	28
Puissance de calcul	25	23	18	31	26	21

Lecture : en 2018, 67 % des sociétés de 250 personnes ou plus implantées en France et payant pour des services de *cloud computing* utilisent des services de courriel en *cloud*.

Champ : sociétés de 10 personnes ou plus des secteurs principalement marchands, hors secteurs agricole, financier et d'assurance, payant pour des services de *cloud computing* en 2018.

Sources : Eurostat ; Insee, enquête TIC-entreprises 2018.

Le *cloud computing* concerne en majorité les sociétés de l'information-communication et des activités spécialisées, scientifiques et techniques

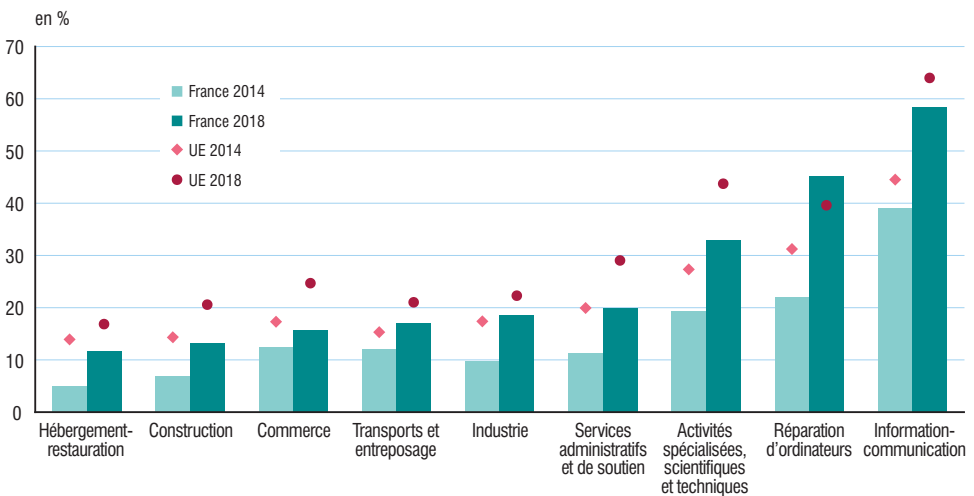
En 2018, en France comme dans l'UE, c'est le secteur de l'hébergement-restauration qui recourt le moins à du *cloud computing* payant (17 % des sociétés dans l'UE, 12 % en France). À l'autre bout du spectre, les sociétés de l'information-communication utilisent le plus le *cloud computing* (64 % des sociétés dans l'UE, 58 % en France), suivies par les sociétés des activités

spécialisées, scientifiques et techniques (44 % des sociétés dans l'UE, 33 % en France), les sociétés de réparation de matériel informatique et les sociétés des activités des services administratifs et de soutien (figure 5).

Bien que les secteurs n'aient pas tous le même recours au *cloud computing*, les différences de structures sectorielles de l'économie entre les pays de l'UE ne suffisent pas à expliquer leurs différences de recours au *cloud computing*. En effet, si tous les pays avaient la même structure sectorielle que la France, pour les sociétés de 10 personnes ou plus, dans l'ensemble, le classement ne serait pas modifié et la position de la France par rapport à l'UE serait identique.

Le type de services que les sociétés achètent dépend beaucoup de leur secteur d'activité : 71 % des sociétés de l'information-communication de l'UE payant du *cloud* recourent à des services de haut niveau, contre 45 % des sociétés industrielles. De plus, parmi les services de haut niveau, l'information-communication est le seul secteur où les sociétés payant du *cloud* le font davantage pour de la puissance de calcul que pour l'exécution de logiciels de comptabilité, en France comme dans l'UE.

5. Recours au *cloud computing* payant selon le secteur d'activité, en France et dans l'UE, entre 2014 et 2018



Note : dans le secteur de la réparation d'ordinateurs, les écarts entre la France et l'UE ne sont pas statistiquement significatifs, en raison du faible nombre de sociétés dans ce secteur.

Lecture : entre 2014 et 2018, la part de sociétés de 10 personnes ou plus de l'hébergement-restauration en France payant pour des services de *cloud* est passée de 5 % à 12 %. Dans le même temps, dans l'UE, elle est passée de 14 % à 17 %.

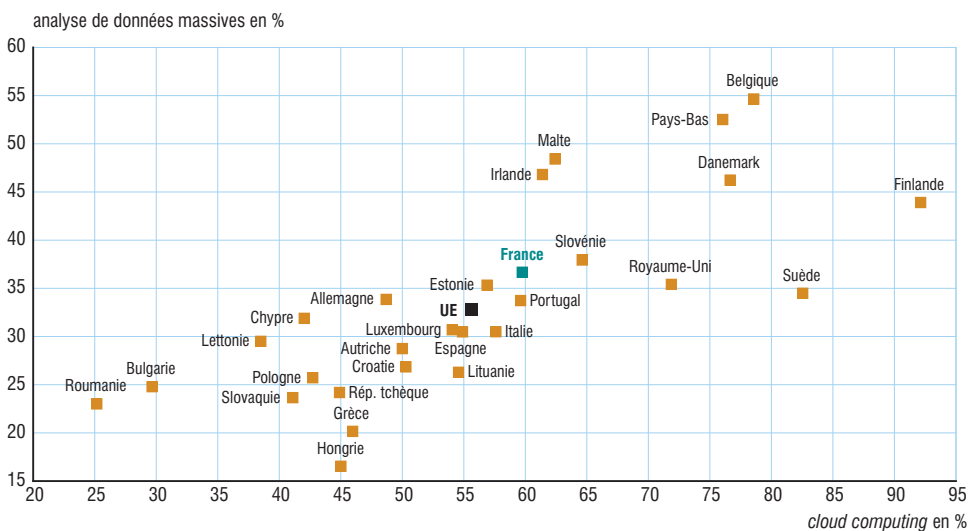
Champ : sociétés de 10 personnes ou plus des secteurs principalement marchands, hors secteurs agricole, financier et d'assurance.

Sources : Eurostat ; Insee, enquête TIC-entreprises 2018.

Cloud computing et analyse de données massives vont de pair pour les grandes sociétés

Dans les pays où le *cloud computing* est une pratique largement adoptée par les grandes sociétés (250 personnes ou plus), l'analyse de **données massives (big data)** est aussi plus répandue au sein de ces grandes sociétés (figure 6). En 2018, dans l'UE, 33 % des sociétés de 250 personnes ou plus ont réalisé ou ont fait réaliser des analyses de données massives. En France, elles sont 37 % et en Belgique et aux Pays-Bas, plus de la moitié.

6. Recours au *cloud computing* et à l'analyse de données massives par les entreprises de 250 personnes ou plus en 2018



Note : pour le Royaume-Uni, faute de données sur l'analyse de données massives pour 2018, c'est la valeur de 2016 qui est utilisée.

En 2018, les sociétés déclarent les analyses de données massives qu'elles ont faites l'année précédente (en 2017).

Lecture : en 2018, 23 % des sociétés de 250 personnes ou plus en Roumanie déclarent avoir réalisé des analyses de données massives l'année précédente et 25 % payent des services de *cloud computing*.

Champ : sociétés de 250 personnes ou plus des secteurs principalement marchands, hors secteurs agricole, financier et d'assurance.

Sources : Eurostat ; Insee, enquête TIC-entreprises 2018.

Cloud computing et *big data* vont de pair pour ces grandes sociétés. En effet, l'analyse de données massives, qui consiste à valoriser des données variées et volumineuses produites en continu par les interactions des systèmes des entreprises, nécessite un outillage informatique spécifique avec une grande capacité d'hébergement de bases de données et une puissance de calcul facilement modulable, en particulier lorsque l'analyse de données massives n'est pas routinière. De plus, lorsque les données massives ne sont pas analysées à l'endroit où elles sont produites, le *cloud* est un moyen de fournir en temps réel un accès distant à ces données. Certains fournisseurs de *cloud computing* proposent des services « clef en main » d'analyse de données massives sur le *cloud*.

Ce lien entre *cloud computing* et *big data* est moins marqué pour les sociétés de plus petite taille. En France, si les petites sociétés sont en retard sur l'usage du *cloud computing* par rapport à l'UE, elles sont en avance pour l'analyse de données massives : 14 % des sociétés de 10 à 49 personnes en France en ont réalisé ou fait réaliser en 2017, contre 10 % dans l'UE. Elles sont au plus 21 % à Malte et entre 17 % et 18 % en Belgique, aux Pays-Bas, en Irlande et en Finlande.

Les données individuelles européennes de l'enquête communautaire TIC-entreprises n'étant pas disponibles, il n'est pas possible de conclure que les sociétés qui utilisent le *cloud computing* sont les mêmes que celles qui analysent des données massives. Une analyse sur les données individuelles françaises permet toutefois de préciser les liens entre ces deux pratiques numériques : à caractéristiques données, les sociétés achetant du *cloud* ont une probabilité plus forte de réaliser des analyses de données massives. Réciproquement, les sociétés réalisant des analyses de données massives achètent plus fréquemment des services de *cloud computing*, en particulier de la puissance de calcul (encadré 4).

Recours au *cloud computing* et analyse de données massives sont liés

En France, une analyse sur les données individuelles est possible pour étudier le profil des sociétés qui recourent au *cloud computing* ou à l'analyse de données massives. En effet, les données européennes sont fournies agrégées par secteur ou par taille, ce qui ne permet pas d'appliquer le même traitement.

Ce recours est étudié toutes choses égales par ailleurs, c'est-à-dire en tenant compte des caractéristiques et pratiques numériques des sociétés en 2018 suivantes : le secteur d'activité, la tranche d'effectif, l'emploi de spécialistes en TIC,

la possession d'un site *web*, le paiement pour de la publicité sur Internet, la réalisation de ventes *web*, les outils utilisables par les employés via un accès mobile et les types de services *cloud* payés. Toutes ces caractéristiques ont un lien avec la probabilité d'acheter du *cloud computing* en 2018 ou d'avoir réalisé des analyses de données massives en 2017 (*figure*). D'autres caractéristiques ont été étudiées, dont le lien avec les pratiques étudiées n'est pas significatif, comme la réalisation de ventes par échange de données informatisé.

Déterminants du recours au *cloud computing* ou à l'analyse de données massives par les sociétés en 2018 en France

	Cloud computing odds ratio	Big data odds ratio
Caractéristiques de la société		
<i>Hébergement-restauration</i>	Réf.	Réf.
Industrie	n.s.	n.s.
Construction	n.s.	1,4
Commerce	n.s.	ns
Transports	n.s.	3,2
Information-communication ; réparation d'ordinateurs	2,5	n.s.
Activités spécialisées scientifiques et techniques	2,1	n.s.
Services administratifs et de soutien ; immobilier	n.s.	n.s.
<i>De 10 à 49 personnes</i>	Réf.	Réf.
De 50 à 249 personnes	1,4	1,3
250 personnes ou plus	2,6	1,6
Pratiques numériques¹		
A analysé des données massives en 2017	1,4	///
Emploie des spécialistes en TIC	2,6	1,6
A un site <i>web</i>	1,6	n.s.
Paye pour de la publicité sur Internet	1,7	1,7
Fait des ventes <i>web</i>	1,4	1,5
Les employés ont un accès mobile au système de courriels de l'entreprise	1,4	n.s.
Les employés ont un accès mobile aux documents de l'entreprise	1,4	1,4
Les employés ont un accès mobile aux applications métiers de l'entreprise	1,4	1,4
Paye de la puissance de calcul en <i>cloud</i>	///	1,8
Paye des services courriels en <i>cloud</i>	///	0,8
Paye l'hébergement de bases de données en <i>cloud</i>	///	1,2
Paye le stockage de fichiers en <i>cloud</i>	///	1,2
Paye des logiciels de gestion de la relation client en <i>cloud</i>	///	1,4

/// : absence de résultat due à la nature des choses. n.s. : non significatif. Réf. : référence. 1. La valeur de référence est l'opposé (la société n'analyse pas de données massives, n'emploie pas de spécialistes en TIC, etc.). Note : un *odds ratio* exprime le rapport de chances entre deux effets, tous les autres effets étant fixés. Les *odds ratio* présentés ici sont significatifs au seuil de 5 %. Les effets /// ne font pas partie du modèle.

Les modèles de régressions logistiques ne permettent pas d'établir de causalité, mais uniquement d'identifier et de quantifier la corrélation entre la variable dépendante (ici le recours au *cloud computing* payant ou la réalisation d'analyses de données massives) et les variables dites « indépendantes ».

Lecture : à caractéristiques et autres pratiques de la société identiques, la probabilité que la société achète du *cloud computing* est multipliée par 1,4 pour les sociétés qui ont réalisé des analyses de données massives en 2017 par rapport à celles qui n'en ont pas analysé.

Champ : sociétés de 10 personnes ou plus implantées en France des secteurs principalement marchands, hors secteurs agricole, financier et d'assurance.

Source : Insee, enquête TIC-entreprises 2018.

Encadré 4 (suite)

Ainsi, le secteur d'activité a peu d'effet propre sur le recours au *cloud*, sauf pour les activités spécialisées, scientifiques et techniques ou l'information-communication, les sociétés de ces secteurs ayant une probabilité au moins deux fois plus forte de recourir au *cloud computing* payant que dans l'hébergement-restauration, à autres caractéristiques fixées. Pour l'analyse de données massives, le secteur des transports a une probabilité trois fois plus forte de recourir aux analyses de données massives que l'hébergement-restauration, toutes choses égales par ailleurs.

Les pratiques numériques des sociétés décrivent plus leur recours au *cloud* ou à l'analyse de données massives que leur secteur. Certaines pratiques sont liées à la fois au recours au *cloud computing*

et à l'analyse de données massives, comme le fait d'employer des spécialistes informatiques, de payer pour de la publicité en ligne, d'effectuer des ventes *web* et de fournir à ses employés un accès aux documents et aux applications de l'entreprise via le réseau mobile.

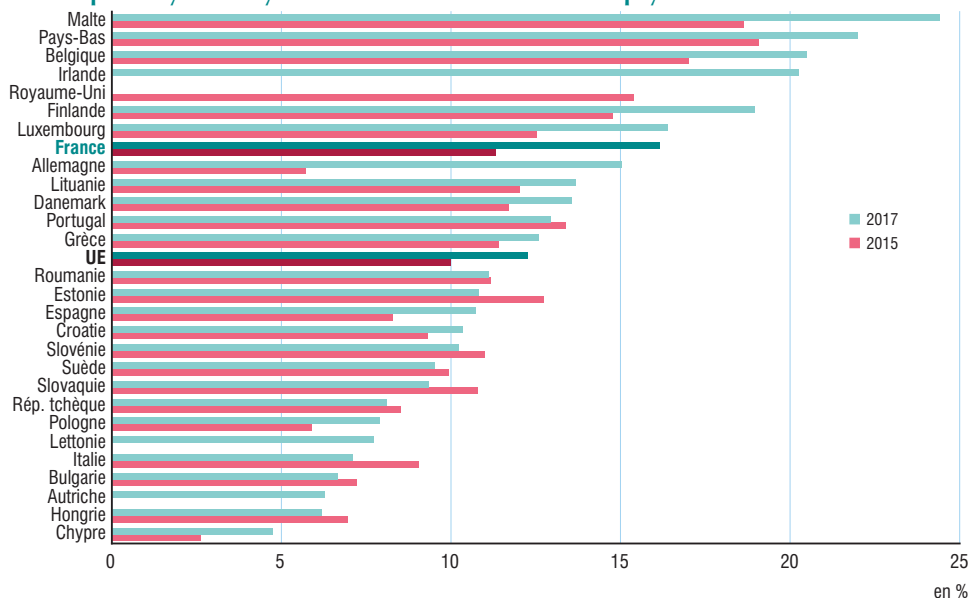
Parmi les services de *cloud computing* payés par les sociétés, toutes choses égales par ailleurs, c'est l'achat de puissance de calcul qui a le lien le plus fort avec le fait de réaliser des analyses de données massives.

Enfin, une fois neutralisés les effets sectoriels ou de pratiques numériques, le recours au *cloud computing* comme aux analyses de *big data* reste d'autant plus fréquent que la taille de la société est grande.

Plus de sociétés analysant des données massives en France que dans l'UE

Dans l'UE, 12 % des sociétés ont analysé ou fait analyser des données massives en 2017, avec peu de changements par rapport à 2015 (10 %). En France, la pratique est plus répandue (16 % en 2017) et a augmenté depuis 2015 (11 %) (figure 7).

7. Entreprises ayant analysé des données massives dans les pays de l'UE en 2015 et en 2017



Note : en Irlande, Lettonie, Autriche, la donnée n'est pas disponible pour 2015. Au Royaume-Uni, la donnée n'est pas disponible pour 2017.

Lecture : en 2017, à Malte, 24 % des sociétés de 10 personnes ou plus ont analysé des données massives. Elles étaient 19 % en 2015.

Champ : sociétés de 10 personnes ou plus des secteurs principalement marchands, hors secteurs agricole, financier et d'assurance.

Sources : Eurostat ; Insee, enquêtes TIC-entreprises 2016 et 2018.

En 2017, les pays où l'analyse de données massives est la plus répandue sont Malte (24 %), les Pays-Bas (22 %), la Belgique et l'Irlande (20 %) ainsi que la Finlande (19 %), comme en 2015. Le Luxembourg et la France suivent à égalité. À l'inverse, en Autriche, Hongrie, Italie, Bulgarie et à Chypre, l'analyse de données massives en 2017 est peu fréquente, et parfois moins répandue que ce qu'elle était en 2015. L'analyse de données massives en 2017 est peu variable selon les secteurs : dans l'UE, elle concerne entre 10 % et 14 % des sociétés de tous les secteurs, à l'exception des transports et entreposage (19 %) et de l'information-communication (27 %). Comme pour le *cloud computing*, les différences sectorielles entre pays ne permettent pas d'expliquer leurs différences de recours à l'analyse de données massives.

En France, plus d'analyses de données de géolocalisation

La plupart du temps, les sociétés qui analysent des données massives analysent au moins un des types de données suivants : des données produites par les appareils et capteurs de l'entreprise elle-même ; des données issues de la géolocalisation d'appareils portables ; des données générées par les médias sociaux. Dans l'UE, l'analyse de données issues d'appareils et capteurs de l'entreprise elle-même concerne 29 % des sociétés ayant réalisé des analyses de données massives en 2017, loin derrière les données de géolocalisation et les données des médias sociaux, analysées par près de la moitié des sociétés réalisant des analyses de données massives.

En France, les sociétés réalisant des analyses de données massives utilisent davantage des données issues de la géolocalisation d'appareils portables (58 % contre 49 % dans l'UE). Cette pratique des sociétés en France est atypique : dans les cinq pays en tête de l'UE (Malte, Pays-Bas, Belgique, Irlande, Finlande), les analyses de données de géolocalisation issues d'appareils portables sont moins courantes qu'en moyenne dans l'UE (28 % à Malte et 32 % en Irlande).

En France, les données issues des médias sociaux ou d'appareils et capteurs sont au contraire moins utilisées que dans les autres pays où l'analyse de données massives est la plus répandue. Ainsi, à Malte, 63 % des sociétés ayant analysé des données massives ont exploité des données issues des médias sociaux et 66 % en Irlande, contre seulement 45 % dans l'UE et 40 % en France.

En France, un tiers des sociétés des transports et de l'entreposage ont mené des analyses de données massives

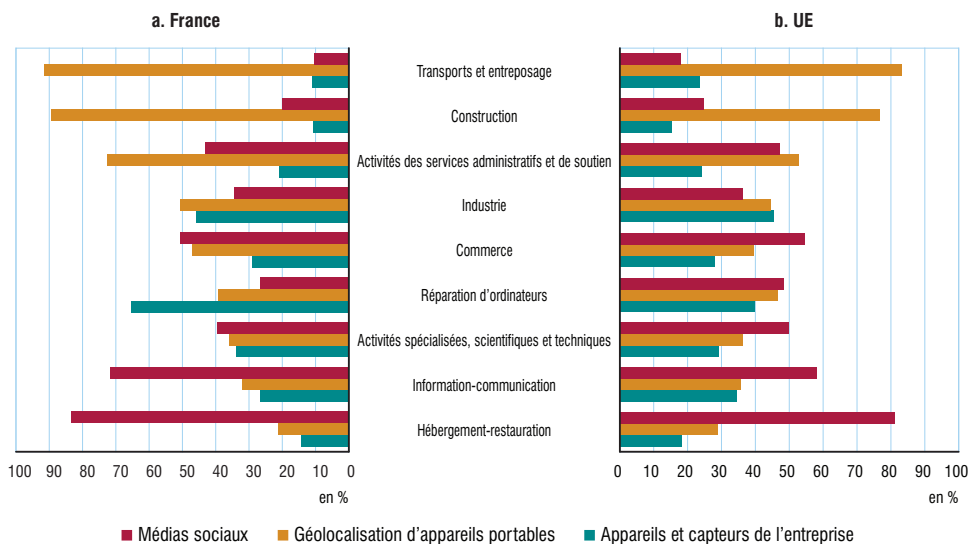
En France, 33 % des sociétés des transports et de l'entreposage ont réalisé en 2017 des analyses de données massives, une part atteinte dans aucun autre pays (au mieux 25 % pour l'Irlande). En 2015, un quart des sociétés de ce secteur en France avaient réalisé des analyses de données massives, à égalité avec le Luxembourg et sans équivalent dans les autres pays.

Les sociétés des transports et de l'entreposage, de la construction et des activités des services administratifs et de soutien sont celles qui utilisent le plus les données de géolocalisation (*figure 8 a* et *b*). L'usage de ces données par ces secteurs est encore plus intense en France que dans l'UE, compensant le fait que les sociétés de France dans l'hébergement-restauration, l'information-communication et la réparation d'ordinateurs recourent relativement moins aux données de géolocalisation que l'ensemble des sociétés de ces secteurs dans l'UE.

Combinés, ces éléments expliquent l'usage globalement plus important des données de géolocalisation en France par rapport aux autres pays où le recours aux analyses de données massives est important.

Les sociétés de l'hébergement-restauration et surtout de l'information-communication sont de grandes consommatrices de données des médias sociaux. Elles les exploitent relativement

8. Type de données massives analysées, par secteur d'activité, en 2017



Lecture : en 2017, 91 % des sociétés du secteur des transports et de l'entreposage en France qui ont réalisé des analyses de données massives ont exploité des données issues de la géolocalisation, contre 83 % pour les sociétés de ce secteur dans l'UE.

Champ : sociétés de 10 personnes ou plus des secteurs principalement marchands, hors secteurs agricole, financier et d'assurance, ayant réalisé des analyses de données massives en 2017.

Sources : Eurostat ; Insee, enquête TIC-entreprises 2018.

plus en France que dans l'UE, au contraire des autres secteurs. En revanche, presque tous les secteurs exploitent moins les données issues de capteurs et d'appareils de l'entreprise en France que dans l'UE.

Les plus grandes sociétés utilisent plus souvent des données issues de capteurs et d'appareils de l'entreprise, en France comme dans l'UE (54 % des sociétés de 250 personnes ou plus qui ont réalisé des analyses de données massives, contre 24 % des sociétés de 10 à 49 personnes). À l'inverse, l'analyse de données de géolocalisation est proportionnellement plus fréquente pour les petites sociétés (51 % de celles qui ont analysé des données massives) que pour les grandes (41 %), cet écart étant accentué en France (62 % versus 43 %). ■

Définitions

Cloud computing

Le *cloud computing* ou *cloud* (informatique en nuage) désigne ici des services informatiques utilisés sur Internet pour accéder à un logiciel, à de la puissance de calcul, à une capacité de stockage, etc. Ces services doivent présenter toutes les caractéristiques suivantes :

- être délivrés par des serveurs informatiques de prestataires ;
- être facilement modulables à la hausse ou à la baisse (par exemple nombre d'utilisateurs ou changement de la capacité de stockage) ;
- une fois installés, pouvoir être utilisés « à la demande », sans interaction humaine avec le fournisseur ;
- être payants, soit par l'utilisateur, soit selon la capacité utilisée, ou être prépayés.

Données massives (*big data*)

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.). Elles présentent les caractéristiques suivantes :

- un volume important issu d'une grande quantité de données générées au cours du temps ;
- une variété concernant les différents formats de données complexes, structurées ou non (par exemple, textes, vidéos, images, voix, documents, données de capteurs, journaux d'activité, 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.

L'analyse de données massives fait référence à l'utilisation de techniques, technologies, algorithmes et logiciels adaptés (*machine learning*, *data mining*, etc.).

Médias sociaux

Les médias sociaux sont des sites et applications *web* qui permettent la création et la publication de contenus générés par l'utilisateur et le développement de réseaux sociaux en ligne en connectant les profils des utilisateurs (*Facebook*, *LinkedIn*, *Twitter*, *YouTube*, *SlideShare*, etc.).

Pour en savoir plus

Cloud computing services used by more than one out of four enterprises in the EU, Eurostat, décembre 2018.

Vacher T., Démoly E., « La timide émergence du *cloud computing* dans les sociétés en 2014 », *Insee Première* n° 1545, avril 2015.

Vacher T., Pradines N., « *Cloud computing*, *big data* : de nouvelles opportunités pour les sociétés », *Insee Première* n° 1643, mars 2017.
