



Bruxelles, le 19.4.2016
COM(2016) 178 final

**COMMUNICATION DE LA COMMISSION AU PARLEMENT EUROPÉEN, AU
CONSEIL, AU COMITÉ ÉCONOMIQUE ET SOCIAL EUROPÉEN ET AU COMITÉ
DES RÉGIONS**

**Initiative européenne sur l'informatique en nuage - Bâtir une économie compétitive des
données et de la connaissance en Europe**

{SWD(2016) 106 final}
{SWD(2016) 107 final}

Introduction

On assiste aujourd'hui à une augmentation spectaculaire de la quantité et de la diversité de données produites. Parallèlement aux données créées par des milliards de personnes utilisant des appareils et services numériques à des fins privées et professionnelles, et aux données générées par les objets connectés, de plus en plus nombreux, d'autres données proviennent de la recherche, de la littérature et des archives numérisées, ainsi que des services publics tels que les hôpitaux et les registres fonciers. Ce phénomène des mégadonnées («big data») crée de nouvelles opportunités pour le partage des connaissances, la recherche et l'élaboration et la mise en œuvre de politiques publiques.

Il devient aussi plus facile d'exploiter ces données grâce à l'informatique en nuage. L'informatique en nuage peut être définie comme la combinaison de trois éléments interdépendants: les infrastructures de données, qui stockent et gèrent les données; les réseaux à haut débit, qui transportent les données; et les ordinateurs de plus en plus puissants qui peuvent être utilisés pour traiter les données. La capacité d'analyse et d'exploitation de ces mégadonnées a des répercussions sur l'économie mondiale et la société, ouvrant la voie à d'importantes innovations industrielles et sociales. Un élément essentiel de répercussions réside dans le changement dans la manière dont est menée la recherche scientifique, au vu de l'évolution rapide vers la [science ouverte](#) que nous connaissons actuellement.

L'informatique en nuage permet de déplacer, de partager et de réutiliser des données sans discontinuité sur les marchés mondiaux et par-delà les frontières, et entre les institutions et les disciplines de recherche. Compte tenu des capacités actuellement disponibles en Europe, les données produites par la recherche et l'industrie de l'UE sont souvent traitées ailleurs et les chercheurs et innovateurs européens ont tendance à se déplacer là où des capacités élevées en matière de données et d'informatique sont plus immédiatement disponibles. Parallèlement, l'Europe étant le plus grand producteur de connaissances scientifiques dans le monde, elle est bien placée pour jouer un rôle moteur au niveau mondial dans le développement d'une informatique en nuage au service de la science.

Si elle veut exploiter pleinement le potentiel de données comme vecteur essentiel de la science ouverte et de la [4^e révolution industrielle](#), l'Europe doit répondre à un certain nombre de questions spécifiques:

- Comment maximiser les incitations au partage des données et accroître la capacité de les exploiter?
- Comment faire en sorte que les données puissent être utilisées aussi largement que possible, dans les différentes disciplines scientifiques et entre les secteurs public et privé?
- Comment mieux interconnecter les infrastructures de données existantes et nouvelles dans l'ensemble de l'Europe?
- Comment coordonner au mieux le soutien mis à la disposition des infrastructures de données européennes dans leur transition vers le calcul à l'échelle exa (exascale)¹?

¹ Le calcul à l'échelle exa désigne les systèmes informatiques d'une capacité d'au moins un exaFLOPS – 10¹⁸ calculs par seconde – soit une vitesse environ 1 000 fois plus grande que les machines actuelles.

Les gains potentiels pour la science, la technologie et l'innovation que promettent les actions visant à relever ces défis ont été soulignés par la [communauté scientifique elle-même](#), mais aussi par les [gouvernements de pays de l'OCDE](#). L'importance que revêt, pour l'ensemble de l'économie et de la société, la capacité à relever ces défis a été confirmée par les États membres de l'UE en 2015². La présente communication propose, en réponse directe à ces défis, une initiative européenne sur l'informatique en nuage à même d'asseoir la place de l'Europe dans l'économie fondée sur les données au niveau mondial³.

L'initiative européenne sur l'informatique en nuage s'appuie sur la stratégie pour un marché unique numérique, qui vise, entre autres, à maximiser le potentiel de croissance de l'économie numérique européenne⁴. Elle vise à développer un environnement fiable et ouvert permettant à la communauté scientifique de stocker, partager et réutiliser des données et résultats scientifiques, le *European Open Science Cloud* (nuage européen pour la science ouverte)⁵. Il s'agit de développer les capacités sous-jacentes de calcul intensif, la connectivité rapide et les solutions d'informatique en nuage à haute capacité dont elle a besoin grâce à une **infrastructure de données européenne**⁶. Axée initialement sur la communauté scientifique, la base d'utilisateurs sera étendue au secteur public et à l'industrie, ainsi qu'à la création de solutions et de technologies qui bénéficieront à tous les secteurs de l'économie et de la société. Pour y parvenir, il faudra consentir un effort de collaboration ouvert à tous ceux qui sont intéressés par l'exploitation de la révolution des données en Europe en tant que composante essentielle de la croissance mondiale.

L'initiative européenne sur l'informatique en nuage s'appuie sur les réalisations de la stratégie européenne en matière d'informatique en nuage⁷ et de la stratégie en matière de calcul à haute performance (CHP)⁸. Elle prolongera des initiatives telles que le projet important d'intérêt européen commun (PIIEC) sur les applications compatibles avec le CHP et les mégadonnées, récemment annoncé⁹. Elle donne une impulsion à la politique définie dans la communication sur les mégadonnées¹⁰ et soutient le programme politique européen en matière de science ouverte, qui vise à améliorer la qualité et l'impact de la science¹¹, en s'appuyant sur les réalisations de l'accès ouvert¹². La présente communication marque le début d'un processus par lequel la Commission s'emploiera, avec les États membres et toutes les parties intéressées, à faire en sorte que l'initiative européenne sur l'informatique en nuage atteigne ses objectifs.

L'initiative européenne sur l'informatique en nuage sera complétée par des mesures supplémentaires au titre de la stratégie pour un marché unique numérique qui couvre les contrats d'informatique en nuage pour les entreprises utilisatrices et le changement de

² Voir les [conclusions](#) du Conseil «Compétitivité», 2015.

³ Discours du Président Juncker, octobre 2015; <http://bit.ly/1Y52pGi>

⁴ COM(2015) 192 final.

⁵ Des travaux préparatoires ont été lancés par l'intermédiaire d'un groupe d'experts de haut niveau de la Commission, chargé d'émettre des conseils sur sa mise en place: <http://bit.ly/1RK7lhh>

⁶ Travaux préparatoires entrepris par l'intermédiaire de groupes consultatifs tels que le groupe de réflexion sur les infrastructures électroniques.

⁷ COM(2012) 529 final et résultats des groupes de travail <http://bit.ly/1QVrvIb>

⁸ COM(2012) 45 final.

⁹ L'objectif est de soutenir le développement de nouvelles utilisations industrielles du CHP et de garantir l'accès à des infrastructures CHP pour la recherche publique et privée, <http://bit.ly/1RMFq0i>

¹⁰ COM(2014) 442 final.

¹¹ Débat d'orientation au sein du Conseil (9385/15); conclusions du Conseil (8970/15).

¹² COM(2012) 401 final.

fournisseur de services d'informatique en nuage, ainsi que par l'initiative en matière de libre circulation des données¹³.

Cinq raisons pour lesquelles l'Europe n'exploite pas encore pleinement le potentiel des données

Premièrement, bon nombre d'entreprises, de communautés de recherche et d'organismes publics européens n'exploitent pas encore pleinement le potentiel des **données** et de leur impact potentiellement transformateur sur les secteurs traditionnels et sur l'organisation de la recherche¹⁴. **Les données provenant de la recherche publique ne sont pas toujours en accès libre**; de même, les données générées ou collectées par les entreprises ne sont bien souvent pas partagées, et pas toujours pour des raisons commerciales. Si certains considèrent encore les données comme un atout à protéger, nombreux sont ceux qui, dans les entreprises (en particulier les PME), les universités et le secteur public, ignorent tout simplement la valeur du partage de données. Cela s'explique notamment par **l'absence de structure claire des mesures d'incitation**, l'absence de récompenses pour le partage de données (principalement dans les milieux universitaires) et de base juridique claire¹⁵ (principalement dans le secteur public) ainsi que par la pénurie de compétences relatives aux données et le manque de reconnaissance de leur valeur (tous secteurs confondus). Le cadre de protection des données de l'UE empêche les restrictions à la libre circulation des données à caractère personnel pour des motifs de protection de la vie privée et des données à caractère personnel. D'autres obstacles juridiques et techniques à la libre circulation des données doivent encore être levés au moyen de la future initiative du marché unique numérique relative à la libre circulation des données.

Deuxièmement, le **manque d'interopérabilité** empêche de relever les grands défis de société qui nécessitent un partage efficient des données et une approche pluridisciplinaire et multi-acteurs, comme, par exemple, le problème du changement climatique, qui ne peut être résolu par les climatologues seuls. Si l'interopérabilité et le partage de données ont déjà fait l'objet de mesures dans certains secteurs (par exemple la localisation des données dans le cadre de la directive INSPIRE¹⁶ et les données en matière de santé dans le cadre de la directive sur les droits des patients¹⁷), de nombreux ensembles de données demeurent néanmoins indisponibles pour les scientifiques, l'industrie, les administrations publiques et les décideurs politiques. Si

¹³ Les propositions législatives éventuelles devront répondre aux exigences de la Commission en matière d'amélioration de la réglementation, conformément aux lignes directrices de la Commission pour une meilleure réglementation (*Better Regulation Guidelines*), SWD(2015) 111.

¹⁴ C'est le cas, par exemple, dans le domaine de la santé (<http://bit.ly/1XEeaTN>) (et des projets du CER BIOTENSORS, DIOCLES, SMAC), de l'astronomie (SparseAstro, p. ex.), du changement climatique, de la migration ou de l'internet (p. ex. DIADEM, MIGRANT, RAPID, THINKBIG).

¹⁵ La directive INSPIRE (2007/2/CE) constitue un acquis pour le partage de données de localisation européennes. Toutefois, le champ d'application de ces dispositions législatives est limité aux données et services spécifiques aux politiques de l'environnement, de lutte contre les catastrophes naturelles et de la santé et les obstacles en matière de politiques de données n'ont pas tous été levés de manière efficace.

¹⁶ Règlement (UE) n°1089/2010 portant modalités d'application de la directive 2007/2/CE

¹⁷ Travaux concernant le réseau de santé en ligne mis en place en vertu de la directive 2011/24/UE relative aux droits des patients, aux infrastructures de services numériques dans le domaine de la santé en ligne, aux services de prescription en ligne et aux services liés aux dossiers de patients, pour échanger des données en matière de santé, et la récente action commune visant à soutenir le rapport sur le réseau de santé en ligne intitulé «Utilisation de l'informatique en nuage dans le domaine de la santé» pour soutenir l'utilisation de données autres que celles concernant des soins directs prodigués à un patient donné.

l'interopérabilité des données administratives nécessite principalement des normes minimales et une sécurité juridique en termes d'accès et d'utilisation, ainsi qu'un soutien pratique¹⁸, le partage des données de la recherche est également entravé par la taille des ensembles de données, ses formats variés, la complexité du logiciel nécessaire pour l'analyser et un cloisonnement bien ancré entre les disciplines. Des «métadonnées» simples¹⁹ permettant d'identifier des données et des spécifications pour le partage de données sont nécessaires pour les rendre largement accessibles et prêtes à être traitées au moyen d'outils communs d'analyse de données à source ouverte. Il convient aussi d'aborder les questions de la protection et de la conservation à long terme des données. Il existe déjà des initiatives de terrain à l'échelle mondiale²⁰ et certains États membres font des progrès dans ce domaine, mais la participation européenne à ces initiatives est limitée et ces efforts sont largement fragmentés.

Troisièmement, la **fragmentation** constitue une entrave à la science fondée sur les données²¹. Les infrastructures de données sont fragmentées entre les domaines scientifiques et économiques, les pays et les modèles de gouvernance. Les politiques d'accès pour la mise en réseau, le stockage de données et les infrastructures informatiques diffèrent. Des infrastructures de données et informatiques déconnectées et lentes freinent la découverte scientifique, génèrent des cloisonnements et ralentissent la circulation des connaissances. Des données de la recherche pouvant être partagées, des outils d'analyse de données ouverts et des systèmes informatiques connectés doivent être disponibles pour la grande majorité des chercheurs²² en Europe, et pas seulement pour les scientifiques de haut niveau appartenant aux disciplines phares des principaux organismes de recherche. De plus, les universités et les centres de recherche européens opèrent généralement dans le cadre de structures nationales et ne disposent pas d'un environnement de dimension européenne pour l'analyse informatique, l'analyse du stockage et l'analyse des données, ce qui rend la coopération scientifique dans l'UE plus difficile, en particulier la coopération pluridisciplinaire basée sur des données²³. Dans le cadre d'une récente consultation publique²⁴, la grande majorité des personnes interrogées a répondu que le nuage européen pour la science ouverte rendrait la science plus efficace en améliorant le partage des ressources aux niveaux national et international.

Quatrièmement, on constate, en Europe, une montée de la demande pour une **infrastructure de calcul à haute performance (CHP) de rang mondial permettant de traiter des données**²⁵ dans les domaines de la science et de l'ingénierie: simulation d'un avion de nouvelle génération complet; modélisation du climat; lien entre génome et santé; compréhension du cerveau humain; tests *in silico* de cosmétiques pour réduire l'expérimentation animale – toutes ces initiatives nécessitent des capacités de calcul à l'échelle exa (exascale). Si, à long terme, l'**informatique quantique** recèle la promesse de résoudre les problèmes de calcul dont l'ampleur dépasse les capacités des superordinateurs

¹⁸ Traité par le programme ISA de la Commission: <http://bit.ly/24DxWUs>.

¹⁹ Il peut s'agir de métadonnées statistiques de haute qualité provenant de statistiques officielles et destinées à renforcer la cherchabilité, l'interopérabilité et l'intégration des données.

²⁰ Plusieurs initiatives mondiales ont été prises dans ce domaine: principes FAIR en matière de données, principes du G8 concernant la science ouverte fondée sur les données issues de la recherche, lignes directrices RDA, recommandations du Forum Belmont, principes de l'OCDE et lignes directrices par discipline.

²¹ La consultation sur l'approche «Science 2.0» a mis en évidence le manque d'intégration des infrastructures existantes comme un obstacle au travail des scientifiques.

²² Les chercheurs ne disposent pas d'un système de stockage et de mise à jour de leurs données (37 %) ou en ignorent l'existence (54 %) (bit.ly/206u6hm).

²³ <http://bit.ly/1SkL9wm>

²⁴ <http://bit.ly/1JEymCY>

²⁵ Demandes de cycles informatiques concernant la double disponibilité de PRACE: <http://bit.ly/1So2sgc>.

actuels²⁶, la compétitivité de l'UE dépend néanmoins aussi du soutien apporté par le CHP aux infrastructures de données paneuropéennes.

À l'échelle mondiale, les États-Unis, la Chine, le Japon, la Russie et l'Inde avancent rapidement. Ces pays ont fait du CHP une priorité stratégique; ils financent des programmes de développement d'écosystèmes CHP nationaux (matériel, logiciels, applications, compétences, services et interconnexions) et travaillent au déploiement de superordinateurs exaflopiques²⁷. L'Europe ne participe pas à la course au CHP à la hauteur de son potentiel économique et de connaissances; elle accuse un retard par rapport à d'autres régions parce qu'elle n'investit pas dans son écosystème CHP et ne tire pas profit des avantages de la propriété intellectuelle dans ce domaine. Du côté de l'offre, l'industrie européenne fournit environ 5 % des ressources CHP dans le monde, alors qu'elle consomme un tiers de celles-ci. L'Europe dépendant de plus en plus d'autres régions pour la technologie critique, elle est exposée à un risque de blocage ou de retard sur le plan technologique ou risque d'être privée de savoir-faire stratégique. L'Europe est également à la traîne en simples termes de puissance de calcul totale: seule une des dix infrastructures de CHP de premier plan est établie dans l'UE, à savoir le Höchstleistungsrechenzentrum Stuttgart, en Allemagne, qui se classe au 8^e rang mondial. Les États-Unis en comptent cinq et la Chine dispose du superordinateur le plus rapide au monde depuis 2013.

Aucun État membre ne dispose à lui seul des ressources financières suffisantes pour développer l'**écosystème CHP nécessaire** dans des délais permettant de rester compétitif par rapport aux États-Unis, au Japon ou à la Chine²⁸. Toutefois, à ce jour, aucune mesure commune n'a été prise pour combler l'écart entre la demande interne et l'offre de l'UE²⁹. L'UE a mis en place un partenariat public-privé contractuel sur le CHP pour développer une technologie exascale, mais il n'existe pas de cadre européen pour l'intégrer dans des systèmes informatiques à grande échelle.

Enfin, les producteurs et les utilisateurs de données scientifiques doivent être en mesure de réutiliser les données et d'utiliser des techniques d'analyse de pointe, telles que la fouille de textes et de données, dans un environnement qui soit au moins aussi fiable que leurs propres installations. Les États membres ont fortement insisté sur l'importance des données de la recherche de l'UE et sur la nécessité de faire en sorte que la science fondée sur les données profite à la société européenne³⁰. Toute utilisation et toute réutilisation de données scientifiques doivent faire en sorte que les données à caractère personnel soient correctement protégées selon les règles de l'UE en matière de protection des données³¹. La présente révision de la législation européenne en matière de droit d'auteur et celle qui suivra³² fournissent des cadres généraux présentant un intérêt dans ce contexte.

²⁶ SWD(2016) 107

²⁷ SWD(2016) 106

²⁸ Le ministère américain de la défense investira 525 millions de dollars pour l'acquisition de 3 systèmes pré-exascale en 2017/2018 («CORAL»). Le Japon prévoit d'investir 1,38 milliard de dollars pour installer un système proche de l'exascale en 2019.

²⁹ Bien que PRACE permette le partage des ressources informatiques de certains États membres, l'acquisition de systèmes CHP est une décision nationale prise en dehors de toute coordination ou de tout financement de l'UE.

³⁰ Conclusions du Conseil (8970/15).

³¹ COM(2012) 9 final.

³² COM(2015) 626 final.

Quelles solutions existent?

1. Nuage européen pour la science ouverte

Le nuage européen pour la science ouverte vise à conférer à l'Europe une position de leader mondial dans le domaine des infrastructures de données scientifiques, afin de permettre aux scientifiques européens de tirer pleinement profit de la science fondée sur les données. Pratiquement, il offrira à 1,7 million de chercheurs européens et à 70 millions de professionnels de la science et de la technologie un environnement virtuel présentant les caractéristiques suivantes: gratuité au point de service, services de stockage, de gestion, d'analyse et de réutilisation des données de la recherche ouverts et sans discontinuité, par-delà les frontières et dans l'ensemble des disciplines scientifiques. Son développement sera porté par la communauté scientifique, qui regroupe les utilisateurs les plus avancés et les plus grands producteurs de travaux scientifiques du monde. Le nuage européen pour la science ouverte sera également ouvert à des fins d'éducation et de formation dans l'enseignement supérieur ainsi que, au fil du temps, aux utilisateurs de l'administration et du monde des entreprises, puisque les technologies développées seront appelées à être utilisées à une plus large échelle.

Le nuage européen pour la science ouverte commencera par fédérer les infrastructures de données scientifiques existantes, aujourd'hui dispersées entre les différents domaines et les différents États membres. L'accès aux données scientifiques sera ainsi plus facile, moins coûteux et plus efficient. Le nuage européen pour la science ouverte permettra de créer de nouveaux débouchés commerciaux et de nouvelles solutions dans des domaines clés tels que la santé, l'environnement ou le transport. Il offrira un environnement sûr où le respect de la vie privée et la protection des données doivent être garantis dès la conception sur la base de normes reconnues, et où les utilisateurs seront protégés contre les risques liés à la sécurité des données et à la responsabilité. Il tirera parti des autres mesures prises par la Commission pour promouvoir la science ouverte en Europe, telles que celles promouvant le libre accès aux publications et données scientifiques dans le cadre d'Horizon 2020, et réunira les principales parties prenantes pour concevoir de concert les prochaines mesures à prendre. La gouvernance du nuage européen pour la science ouverte sera déterminée à l'issue d'un processus complet d'élaboration qui est déjà en cours.

Plus précisément, pour développer le nuage européen pour la science ouverte, il conviendra:

- **de faire en sorte que toutes les données scientifiques produites par le programme Horizon 2020 soient ouvertes par défaut.** Cela permettra d'étendre le projet pilote actuel³³, dans le cadre duquel les projets concernés mettent en œuvre des plans de

³³ Le projet pilote sur le libre accès aux données de la recherche d'Horizon 2020 couvre: les technologies futures et émergentes, les infrastructures de recherche, les technologies de l'information et de la communication, les thèmes liés à la nanosécurité et à la modélisation dans les nanotechnologies, les matériaux de pointe, les systèmes de fabrication et de transformation avancés et le programme de travail sur la biotechnologie, et des thèmes sélectionnés dans le cadre du pilier «défis de société», tels que la sécurité alimentaire, l'agriculture et la sylviculture durables, la recherche marine, maritime et dans le domaine des eaux intérieures, ainsi que la bioéconomie; la lutte contre le changement climatique, l'environnement, l'utilisation efficace des ressources et matières premières - à l'exception des matières premières; l'Europe dans un monde en évolution – des sociétés ouvertes à tous, innovantes et capables de réflexion; la science avec et pour la société ainsi que l'activité transversale et le domaine d'intérêt «Villes intelligentes et durables». Il est à noter que les projets qui ne font pas partie de ces «domaines essentiels» peuvent encore y être intégrés sur une base volontaire.

gestion des données visant à rendre les données de la recherche faciles à trouver, accessibles, interopérables et réutilisables (principes FAIR)³⁴;

- **de sensibiliser à cette question et de changer les structures d'incitation** des universités, de l'industrie et des services publics à partager leurs données, et d'améliorer la formation, la maîtrise des notions et les compétences en matière de bonne gestion des données. Parallèlement, les principes et lignes directrices relatifs à l'accès aux données de la recherche en Europe³⁵ seront réexaminés de manière à renforcer et à coordonner leur mise en œuvre;
- d'élaborer des spécifications concernant **l'interopérabilité et le partage des données** entre disciplines et infrastructures, en s'appuyant sur les initiatives existantes telles que l'alliance des données de la recherche et le Forum Belmont, ainsi que des dispositions légales telles qu'INSPIRE. Les besoins en matière de normalisation qui apparaîtront au fil du temps seront pris en compte par le biais des priorités du marché unique numérique en matière de normalisation dans le domaine des TIC;
- de créer une **structure de gouvernance paneuropéenne adaptée aux objectifs poursuivis** de manière à fédérer les infrastructures de données scientifiques et à remédier au problème de leur fragmentation. Le dispositif institutionnel supervisera le financement à long terme, la durabilité, la conservation et la gestion des données. Il s'appuiera sur les structures existantes de manière à associer les utilisateurs scientifiques, les organismes de financement de la recherche et les responsables de la mise en œuvre³⁶;
- de développer les **services d'informatique en nuage pour la science ouverte**. Soutenus par l'infrastructure de données européenne, ils permettront aux chercheurs de trouver des données de recherche partagées et d'y accéder, d'utiliser des logiciels analytiques de pointe et des ressources de calcul à haute performance et de prendre connaissance des meilleures pratiques scientifiques fondées sur les données issues des disciplines de pointe;
- **d'élargir la base d'utilisateurs scientifiques** du nuage européen pour la science ouverte aux chercheurs et innovateurs de toutes les disciplines et de tous les États membres, ainsi que de pays partenaires et d'initiatives mondiales, afin qu'ils contribuent à l'excellence et tirent profit de l'initiative³⁷.

L'initiative renforcera d'autres actions en matière de science ouverte que le Conseil³⁸ et le Parlement européen³⁹ ont appelées de leurs vœux, ainsi que les actions à entreprendre dans le cadre du prochain programme stratégique en matière de science ouverte de la Commission. Elle encouragera le recours aux bonnes pratiques en matière de cherchabilité et d'accessibilité des données et aidera les chercheurs à faire reconnaître et récompenser leurs compétences en matière de données; elle facilitera la reproductibilité des résultats et limitera le gaspillage de

³⁴ Les clauses de non-participation existantes, dans le cadre desquelles le libre accès aux données serait contraire à une application commerciale future ou à la protection des données privées et des données à caractère personnel, à la sécurité et à la protection des informations classifiées de l'UE, seront maintenues. L'analyse des projets pilotes a montré que la plupart de ces derniers appliquaient une politique d'ouverture des données, mais que les clauses de non-participation étaient également importantes.

³⁵ C(2012) 4890 final.

³⁶ ESFRI, INSPIRE, eIRG, GEANT, PRACE, ELIXIR, le Forum Belmont et initiatives fédératrices analogues.

³⁷ Les nouvelles initiatives de la Commission peuvent être financées au titre des Fonds structurels et d'investissement européens, à condition que les États membres conviennent de les financer et d'adapter leurs programmes opérationnels en conséquence.

³⁸ Conclusions du Conseil (8970/15).

³⁹ Rapport du Parlement européen 2015/2147(INI).

données telles que les données d'essais cliniques (intégrité de la recherche); elle contribuera à la clarification du modèle de financement pour la production et la conservation de données, réduisant ainsi les activités d'acquisition de rentes et amorçant le marché des services de recherche innovants (fouille de textes et de données avancée, par exemple). L'initiative peut également aider à régler les problèmes de compensation de données («data clearance») et de protection des données à caractère personnel⁴⁰. La Commission consultera les acteurs concernés et travaillera avec les fournisseurs R&D sur la nécessité de mettre en œuvre des lignes directrices pour le domaine scientifique en ce qui concerne la politique de l'Union et la législation de l'Union en matière de protection des données, et sur la nécessité de veiller à ce que l'initiative mette en œuvre les principes juridiques «dès la conception», c'est-à-dire au stade le plus précoce possible.

Actions	Calendrier
La Commission collaborera avec les partenaires politiques et de recherche au niveau mondial afin de favoriser la coopération et de créer des conditions identiques pour tous en matière de partage de données scientifiques et de science fondée sur les données.	À partir de 2016.
La Commission utilisera les programmes de travail d'Horizon 2020 afin de fournir un financement pour intégrer et consolider les plateformes d'infrastructures électroniques, de fédérer les infrastructures de recherche existantes et les nuages scientifiques et de soutenir le développement de services en nuage pour la science ouverte.	À partir de 2016.
La Commission fera du libre accès aux données de recherche l'option par défaut, tout en garantissant des clauses de non-participation, pour tous les nouveaux projets du programme Horizon 2020.	À partir de 2017.
La Commission procédera à un réexamen de la recommandation de la Commission de 2012 relative à l'accès aux informations scientifiques et à leur conservation ⁴¹ en vue d'encourager le partage des données scientifiques et la création de mécanismes d'incitation, de systèmes de récompenses et de programmes d'enseignement et de formation permettant aux chercheurs et aux entreprises de partager les données, en étroite relation avec l'initiative du marché unique numérique relative à la libre circulation des données.	À partir de 2017.
La Commission collaborera avec les États membres pour relier les infrastructures européennes de recherche prioritaires ⁴² au nuage européen pour la science ouverte.	À partir de 2017.

⁴⁰ Tout en respectant pleinement les articles 7 et 8 de la Charte des droits fondamentaux de l'Union européenne et les dispositions actuelles et à venir sur l'utilisation des données à des fins de recherche, l'initiative peut développer, par exemple, des services de fouille de textes et de données respectueux des DPR, de contrôle d'accès pour différents usages, d'anonymisation irréversible de données sensibles avant la fusion de données et d'«espaces de données à caractère personnel» afin de préserver la vie privée et de stimuler le recours à des usages innovants, ou s'appuyer sur l'octroi de licences lisibles par machine et des métadonnées liées à la vie privée jointes à des ensembles de données accessibles via l'informatique en nuage, et fournir des lignes directrices et des bonnes pratiques pour les processus organisationnels conformes soutenant l'initiative. Il s'agit certes là d'outils et processus techniques, dès la conception et par défaut, mais ils peuvent contribuer à réduire l'apparition de mauvaises pratiques et la non-conformité avec les dispositions juridiques.

⁴¹ C(2012) 4890 final.

⁴² Telles que recensées par le Forum stratégique européen pour les infrastructures de recherche (ESFRI) (<http://bit.ly/1pfqOe7>).

En collaboration avec les parties prenantes et les initiatives mondiales en la matière, la Commission œuvrera à la mise en place d'un plan d'action pour l'interopérabilité des données scientifiques, et notamment les métadonnées, les spécifications et la certification.	Fin 2017 au plus tard
--	-----------------------

2. Infrastructure de données européenne

L'**infrastructure de données européenne**, une fois pleinement mise en œuvre, servira de fondement au nuage européen pour la science ouverte. L'Europe a besoin de capacités CHP intégrées de niveau mondial, d'une connectivité à haut débit et de services de données et logiciels de pointe⁴³ pour ses scientifiques et pour d'autres utilisateurs clés de l'industrie (y compris les PME) et du secteur public. Cette infrastructure permettra de tirer pleinement parti de la valeur des mégadonnées et du numérique par défaut⁴⁴. L'infrastructure de données européenne aidera également l'UE à se classer parmi les principales puissances mondiales en matière de calcul intensif en **produisant, vers 2022, des superordinateurs exascale basés sur la technologie de l'UE, ce qui la placerait à l'une des 3 premières places à l'échelon mondial**. L'Europe devrait s'efforcer de disposer d'au moins deux sources de cette technologie.

La stratégie CHP existante⁴⁵ soutient certes la recherche et le développement de technologies CHP commercialisables, mais ne prévoit pas la réalisation d'un superordinateur exascale. L'infrastructure de données européenne rassemblera les ressources et capacités nécessaires afin de fermer la chaîne qui va de la recherche et du développement à la réalisation et l'exploitation des systèmes CHP exascale conçus conjointement avec les utilisateurs et les fournisseurs. Il s'agit notamment de garantir la connectivité des données et le stockage de mégadonnées de manière à assurer la disponibilité des services de calcul intensif dans l'ensemble de l'UE, indépendamment de l'endroit où se trouvent les superordinateurs. Un premier pas a récemment été franchi dans ce sens par le Luxembourg, la France, l'Italie et l'Espagne, avec un projet **important d'intérêt européen commun (PIIEC) sur les applications compatibles avec le CHP et les mégadonnées**⁴⁶.

S'appuyant sur les infrastructures et services paneuropéens de calcul à haute performance (PRACE), le réseau à grande vitesse transeuropéen (GÉANT), le partenariat public-privé contractuel sur le CHP⁴⁷, l'entreprise commune ECSEL⁴⁸ et le PIIEC sur les applications compatibles avec le CHP et les mégadonnées, la Commission et les États membres participants:

- favoriseront un écosystème CHP capable de développer une nouvelle technologie

⁴³ Notamment les services existants d'OpenAIRE, EUDAT, EGI, IndigoDataCloud, HelixNebula, PRACE et GÉANT.

⁴⁴ Le «numérique par défaut» désigne les services et processus qui sont rendus accessibles en ligne ou sous une forme numérique par défaut.

⁴⁵ COM(2012) 45 final.

⁴⁷ <http://bit.ly/1QxERan>

⁴⁷ <http://bit.ly/1WZH8wF>

⁴⁸ <http://www.ecsel-ju.eu>

européenne telle que celle des **puces CHP de faible puissance**⁴⁹;

- intégreront des technologies dans des systèmes prototypes par la conception conjointe⁵⁰ de solutions et l'acquisition de systèmes CHP; l'infrastructure CHP qui en résultera mettra l'accent sur les **superordinateurs de capacité haut de gamme** reliés à des centres de calcul nationaux de l'UE de taille moyenne et à des infrastructures de données et de logiciels paneuropéennes pour offrir un service de calcul intensif;
- fourniront une **connectivité sans discontinuité, à haut débit, fiable et sûre** de manière à rendre le CHP accessible dans l'ensemble de l'UE; le réseau à grande vitesse transeuropéen (GÉANT) et les réseaux nationaux pour la recherche et l'éducation (NREN) relient déjà entre eux 50 millions de chercheurs et d'étudiants; ces infrastructures seront modernisées pour faire face à l'augmentation des volumes de données à transférer et à l'extension de la base d'utilisateurs.

L'**infrastructure de données européenne** contribuera à la numérisation de l'industrie tout en développant des plateformes européennes pour de nouvelles applications stratégiques (recherche médicale, industrie aérospatiale, énergie, par exemple) et en favorisant l'innovation industrielle. Elle **élargira la base d'utilisateurs du CHP** en offrant un accès plus facile, grâce à l'informatique en nuage, aussi bien pour les chercheurs dans les principales disciplines scientifiques que pour la «longue traîne» de la recherche. L'industrie, en particulier les PME dépourvues de capacités internes, et les autorités publiques (par exemple, les villes intelligentes et le transport) profiteront de ressources, applications et outils d'analyse CHP en nuage et faciles à utiliser⁵¹. Dans ce contexte, la Commission favorisera le déploiement de capacités de traitement et d'exploitation des données provenant des satellites Sentinelle, des informations provenant des services de Copernicus et d'autres données d'observation de la Terre, de manière à permettre l'enrichissement mutuel entre les différents ensembles de données, à encourager le développement de produits et services innovants et à optimiser les avantages socio-économiques des données d'observation de la Terre en Europe.

L'infrastructure de données européenne opérera en collaboration avec les centres de données scientifiques et publiques nationaux et régionaux. Elle développera et mettra en œuvre les bonnes pratiques basées sur des systèmes de certification et des normes et spécifications européennes et mondiales communes⁵² pour remédier au manque actuel d'interopérabilité entre les centres de données nationaux et disciplinaires⁵³.

⁴⁹ Des machines exascale économes en énergie auraient une incidence sur l'éventail complet des applications de calcul et procureraient à l'Europe des avantages techniques, économiques et sociaux. À l'heure actuelle, une seule machine exascale aurait besoin, pour fonctionner, d'une centrale électrique dédiée d'une puissance de 700 MW, ce qui suffirait pour alimenter en électricité 140 000 ménages pendant un an. Des puces de faible puissance sont donc nécessaires.

⁵⁰ La conception conjointe est une approche de la conception visant à impliquer activement les clients et utilisateurs dans le processus de conception de manière à faire en sorte que le résultat réponde à leurs besoins et soit utilisable.

⁵¹ <http://bit.ly/1pqny20>

⁵² RDA-Europe a lancé l'interaction avec le groupe multipartite sur la normalisation dans le domaine des TIC pour présenter la mise en œuvre des meilleures pratiques pour l'interopérabilité des infrastructures de données élaborées dans le cadre de l'alliance des données de la recherche (Research Data Alliance).

⁵³ Notamment les spécifications relatives aux données et services géographiques interopérables (INSPIRE).

L'infrastructure de données européenne comportera une structure de gouvernance pour la gestion et le développement de l'infrastructure et des services de données⁵⁴, la prise de décision sur le financement, la viabilité à long terme et la sécurité. La gouvernance devrait impliquer les utilisateurs (le nuage européen pour la science ouverte et d'autres utilisateurs à long terme tels que le secteur public), les responsables de la mise en œuvre (PRACE, GEANT) et les bailleurs de fonds, et s'appuyer sur les structures de gouvernance existantes.

Actions	Calendrier
<p>La Commission et les États membres participants doivent élaborer et déployer une infrastructure CHP, de données et de réseau européenne à grande échelle, notamment par les initiatives suivantes:</p> <ul style="list-style-type: none"> - acquisition de deux prototypes de superordinateurs exascale de conception conjointe et de deux systèmes opérationnels qui se classeront parmi les trois premiers du monde; - création d'un centre de mégadonnées européen⁵⁵, - modernisation du réseau général de recherche et d'innovation (GEANT) et intégration des réseaux de services publics européens. 	<p>2016-2020</p> <p>À partir de 2018.</p> <p>À partir de 2016.</p> <p>À partir de 2016.</p>

Exploiter le potentiel des technologies quantiques

La prochaine percée en matière de calcul intensif et de mise en réseau sûre peut s'appuyer sur les technologies quantiques. Les entreprises de premier plan en Europe, Asie-Pacifique et Amérique du Nord commencent à investir dans les technologies quantiques, mais un niveau d'investissement plus élevé est nécessaire pour mettre au point des produits commercialisables. L'Europe doit être à la pointe de ces futures avancées⁵⁶. L'infrastructure de données européenne doit être complétée par une initiative phare ambitieuse, à long terme et à grande échelle pour exploiter pleinement le potentiel des technologies quantiques, accélérer leur développement et proposer des produits commerciaux aux utilisateurs publics et privés. La Commission européenne entamera les préparatifs du lancement de l'initiative phare, notamment la consultation des parties intéressées et l'analyse d'impact, en tenant compte des résultats de l'évaluation intermédiaire du programme Horizon 2020 d'ici la fin de 2017⁵⁷.

Action	Calendrier
<p>La Commission européenne entamera les préparatifs du lancement de l'initiative phare, notamment la consultation des parties intéressées et l'analyse d'impact⁵⁸, en tenant compte des résultats de l'évaluation</p>	<p>2016-2019</p>

⁵⁴ Sur la base des services existants d'OpenAIRE, EUDAT, EGI, IndigoDataCloud, HelixNebula, PRACE et GÉANT.

⁵⁵ Hébergé par exemple par le CCR pour les données pluridisciplinaires mais axé sur les données géographiques d'INSPIRE/GEOSS/Copernicus.

⁵⁶ <https://goo.gl/zBVi8N>

⁵⁷ SWD(2016) 107

⁵⁸ L'analyse d'impact s'inscrira dans le cadre du processus d'élaboration des programmes de financement concernés dans la perspective financière d'après 2020. Toute mesure de mise en œuvre supplémentaire susceptible d'avoir une incidence importante peut nécessiter des analyses d'impact distinctes et individuelles.

intermédiaire du programme Horizon 2020 d'ici la fin de 2017 ⁵⁹ , avec l'ambition de lancer la phase de montée en puissance en 2018 ⁶⁰ .	
--	--

3. Élargissement de l'accès et instauration d'un climat de confiance

Le recours aux services d'informatique en nuage dans le secteur public est inégal et lent⁶¹. Cela s'explique par le manque de confiance et le peu de synergies existant entre le secteur public et les universités. La fragmentation en termes d'infrastructures de données constitue un obstacle à la constitution d'une masse critique et à la mise au point de solutions communes pour les différents groupes d'utilisateurs. La **base d'utilisateurs du nuage européen pour la science ouverte et de l'infrastructure de données européenne sera élargie au secteur public**, par exemple au moyen de projets pilotes à grande échelle impliquant l'administration en ligne⁶² et les parties prenantes du secteur public et en ouvrant progressivement l'infrastructure de données européenne aux **utilisateurs de l'industrie** et du secteur public afin d'atteindre une dimension européenne. Au fil du temps, le nuage européen pour la science ouverte garantira la visibilité, l'accessibilité et l'exploitabilité intégrales des données publiques pour les scientifiques, les décideurs politiques et les entreprises. Les enseignements tirés fourniront des orientations concrètes pour l'adoption de services en nuage par les administrations publiques dans toute l'Europe.

Le secteur public générant des quantités colossales de données (système d'observation de la Terre Copernicus, données de localisation INSPIRE, par exemple) et devant disposer d'une plus grande capacité informatique (par exemple pour les systèmes d'information en temps réel sur le trafic et les itinéraires, les applications en matière de ville intelligente ou la modélisation des politiques), il tirera profit des économies d'échelle, de la flexibilité et de la continuité. Le public bénéficiera donc de services publics moins coûteux, plus rapides, de meilleure qualité et interconnectés et de l'amélioration de l'élaboration des politiques sur la base de services abordables et sûrs à forte intensité de calculs et de données.

De la même manière, le nuage européen pour la science ouverte et l'infrastructure de données européenne bénéficieront aux entreprises, et notamment aux PME, qui ne disposent pas d'un accès facile et financièrement abordable au stockage et aux services de données et à l'informatique de pointe. Des mesures seront prises pour élargir progressivement la base d'utilisateurs aux PME innovantes et à l'industrie, par l'intermédiaire de centres d'excellence en matière de données et de logiciels et de pôles d'innovation en matière de services de données pour les PME. Ces mesures nécessiteront une étroite coopération avec le secteur privé: PME, grands utilisateurs scientifiques et industriels du CHP, et secteur des services en nuage, qui doit y être associé dès le début.

En outre, l'initiative européenne sur l'informatique en nuage devra respecter des **normes élevées en matière de qualité, de fiabilité et de confidentialité**, afin de garantir la protection des données à caractère personnel et de la propriété intellectuelle, **ainsi qu'en matière de sécurité** - en termes de résilience et de protection contre l'intrusion. Les mécanismes existants relevant du secteur public - notamment le mécanisme pour l'interconnexion en Europe (MIE)

⁵⁹ SWD(2016) 107

⁶⁰ Initiatives-phares FET décrites dans les documents de référence Horizon 2020.

⁶¹ SMART 2013/0043: les organismes du secteur public sont distancés par le secteur privé, avec une différence de 10 %, en 2013, dans l'utilisation des services d'informatique en nuage.

⁶² Plan d'action 2016-2020 de l'UE en matière d'administration en ligne: accélérer la transformation numérique du secteur public.

et les composantes de l'infrastructure de services numériques (ISN) liées à la confiance et à la sécurité - peuvent être réutilisés et déployés par la communauté scientifique à des fins d'économies de coûts, de facilité d'accès et de cohérence globale. Le cadre général sera constitué par les règles générales en matière de protection des données, la directive SRI⁶³ et la révision de la législation européenne en matière de droit d'auteur. Compte tenu de la dimension mondiale de l'informatique en nuage, il est essentiel que l'économie européenne des données reste connectée au reste du monde et que les normes mondiales en matière de protection des données soient renforcées pour atteindre un niveau élevé globalement équivalent à celui qui prévaut en Europe.

L'élaboration de normes appropriées s'inscrit dans le cadre des priorités du marché unique numérique concernant le plan pour la normalisation des TIC⁶⁴; un système de certification approprié sera conçu à l'échelon de l'UE pour garantir la sécurité, la portabilité des données et l'interopérabilité dans le respect des prescriptions légales⁶⁵, notamment le système de certification déjà prévu par le règlement général sur la protection des données pour garantir la sécurité des données personnelles. Il existe certes un certain nombre de systèmes de certification⁶⁶, mais leur champ d'application et leur application varient considérablement, et il n'existe pas d'approche commune concernant les exigences minimales pour l'acquisition ou la gestion de ressources d'informatique en nuage du secteur public. À cet égard, la collaboration avec les entreprises et les autorités publiques mettra les capacités des premières en adéquation avec les exigences de la science et du secteur public.

L'élargissement de l'accès au nuage européen pour la science ouverte et à l'infrastructure de données européenne sera réalisé conformément à la législation appropriée, notamment pour ce qui est de la réutilisation des données à d'autres fins.

Actions	Calendrier
<p>En partenariat avec les entreprises et le secteur public, la Commission s'engage à:</p> <ul style="list-style-type: none"> - adapter les solutions CHP et de mégadonnées à un environnement en nuage afin de permettre un large accès, notamment pour les PME; - développer un écosystème pour renforcer le secteur de l'informatique en nuage en Europe, en utilisant le nuage européen pour la science ouverte comme banc d'essai pour la mise au point de solutions technologiques innovantes en matière d'informatique en nuage; - créer une plateforme permettant aux pouvoirs publics d'ouvrir leurs données et services, de manière à créer une base de type «administration en tant que service» (GaaS) pour l'UE. 	2016-2020
<p>Afin de faciliter le recours aux technologies fondées sur les mégadonnées, la Commission fournira un environnement d'essai pour les mégadonnées (projets pilotes à grande échelle) pour les administrations publiques, notamment dans le cadre du PIIEC proposé.</p>	À partir de 2016.

⁶³ COM (2013) 48

⁶⁴ COM(2016) 176

⁶⁵ Règlement (CE) n° 765/2008

⁶⁶ <https://resilience.enisa.europa.eu/cloud-computing-certification>

En collaboration avec les entreprises et les États membres, la Commission favorisera l'utilisation des certificats et normes existants en la matière et – le cas échéant – la création de systèmes de certification et de labellisation au niveau européen, en particulier pour soutenir les marchés publics de services en nuage.	À partir de 2016.
--	-------------------

Incidences financières

La transformation numérique en Europe constitue une entreprise à grande échelle. Différentes sources de financement de l'UE peuvent être recensées pour l'initiative européenne sur l'informatique en nuage:

- Programme-cadre pour la recherche et l'innovation «Horizon 2020» (Horizon 2020)
- Mécanisme pour l'interconnexion en Europe (MIE)
- Fonds structurels et d'investissement européens
- Fonds européen pour les investissements stratégiques

Différentes sources de financement sont nécessaires pour soutenir le cycle d'investissement dans son intégralité. Les grands projets d'infrastructure sont soutenus, au départ, par des subventions publiques et, à mesure qu'ils gagnent en maturité, par des instruments de partage des risques et des instruments fondés sur le marché. Toutefois, de telles initiatives nécessitant des efforts cohérents et coordonnés, la fragmentation des sources budgétaires disponibles constitue clairement un inconvénient.

Le financement existant au titre d'Horizon 2020 permettra de soutenir le nuage européen pour la science ouverte et de donner un coup de fouet à l'infrastructure de données européenne. Le montant des investissements publics et privés supplémentaires requis s'élevait initialement à 4,7 milliards d'euros pour la période de 5 ans visée. Ce montant se décompose comme suit: 3,5 milliards d'euros pour les infrastructures de données⁶⁷, 1 milliard d'euros pour une initiative phare de grande envergure à l'échelle de l'UE dans le domaine des technologies quantiques et 0,2 milliard d'euros pour les actions relatives à l'élargissement de l'accès et au renforcement de la confiance. Des dispositions supplémentaires seront examinées avec les États membres pour élargir le soutien au nuage européen pour la science ouverte au-delà d'Horizon 2020. Au fil du temps, cette initiative générera des recettes par elle-même à mesure que la communauté scientifique, les jeunes pousses innovantes et le secteur public y recourront.

La Commission a l'intention de proposer des moyens de combiner les différentes sources de financement à l'échelon de l'UE et au niveau national pour réaliser intégralement les objectifs de la présente communication; elle les examinera avec les États membres après une évaluation, des analyses d'impact et une consultation appropriées. Des infrastructures aussi ambitieuses nécessiteront un engagement fort de la part des États membres, notamment par la mobilisation des fonds structurels et des garanties du Fonds européen pour les investissements stratégiques⁶⁸, mais également par des investissements importants du secteur privé et des mécanismes de coordination appropriés. À cet égard, la proposition de projet important

⁶⁷ SWD(2016) 106

⁶⁸ Les services de conseil de la BEI au titre de la plateforme européenne de conseil en investissement seront également mis à contribution.

d'intérêt européen commun (PIIEC) sur le CHP) et les mégadonnées montre les possibilités et les effets positifs de l'engagement des États membres.

Actions	Calendrier
En coopération avec les États membres et les parties prenantes, la Commission examinera les mécanismes de gouvernance et de financement appropriés pour le nuage européen pour la science ouverte et l'infrastructure de données européenne et définira une feuille de route pour la mise en œuvre.	À partir de 2016.
La Commission présentera des approches pour combiner les différents flux de financement en vue d'une discussion avec les États membres et les acteurs concernés, afin de réaliser les objectifs de la présente communication.	2016

CONCLUSIONS

L'initiative européenne sur l'informatique en nuage vise à aider la science, l'industrie et les pouvoirs publics européens à accéder à des infrastructures de données et à des services en nuage d'envergure internationale dans la mesure où ils deviennent des facteurs de réussite déterminants dans l'économie numérique.

Une initiative européenne sur l'informatique en nuage devrait offrir à tout centre de recherche, à tout projet de recherche et à tout chercheur en Europe les capacités de calcul intensif et de stockage et d'analyse des données d'envergure mondiale dont ils ont besoin pour réussir à s'intégrer dans le système d'innovation mondial fondé sur les données.

L'initiative permettra d'élargir la base d'utilisateurs des infrastructures et services au secteur public et à l'industrie, y compris aux PME, en garantissant un niveau adéquat de sécurité, de portabilité des données, d'interopérabilité et de conformité aux exigences juridiques européennes.

La mesure dans laquelle les États membres et le secteur privé saisissent les bénéfices à escompter en relevant ce défi et s'engagent à unir leurs efforts pour y arriver déterminera la réussite de l'initiative.