

COMMUNICATION DE LA COMMISSION EUROPÉENNE

# INFRASTRUCTURES TIC POUR LA SCIENCE EN LIGNE



Bruxelles, 5.3.2009  
COM(2009) 108 final



Commission européenne  
Société de l'information et médias

## **NOTICE LÉGALE**

Par la Commission des Communautés européennes, direction générale de la Société de l'information et médias.

La Commission européenne ou toute personne agissant en son nom ne sont pas responsables de l'utilisation qui pourrait être faite de l'information contenue dans la présente publication. La Commission européenne n'est pas responsable des sites web externes auxquels il est fait référence dans la présente publication.

Les opinions développées dans la présente publication sont celles des auteurs et ne reflètent pas nécessairement la position de la Commission européenne.

Office des publications de l'Union européenne- Luxembourg, 2009

ISBN 978-92-79-11285-0

doi 10.2759/42447

© Communautés européennes, mai 2009

Reproduction autorisée, moyennant mention de la source.

# Avant-Propos

Nous sommes confrontés à une crise économique sans précédent qui domine l’agenda politique. En même temps, nous devons répondre à des défis globaux tels que le changement climatique, la préservation de la biodiversité, l’énergie durable et la santé publique.

Certains s’inquiètent de ces menaces et de ces défis. Mais, ne serait-il pas plus judicieux de les transformer en opportunités?

De mon point de vue, l’Union Européenne dispose du potentiel pour réagir à la crise en continuant d’intensifier son effort pour tirer le meilleur des connaissances et de la puissance d’innovation en Europe. Nous devons nous efforcer de renforcer notre avance en matière scientifique ainsi que réaffirmer le rôle de l’Europe comme partenaire global compétitif et comme centre d’excellence.



Cet effort requiert, en particulier, le déploiement d’infrastructures de technologies de l’information et de communication (dénommées e-Infrastructures) pouvant rassembler les scientifiques par delà les distances, et leur donner accès à des données ou des moyens de calcul uniques. C’est ce que les e-Infrastructures, comme le réseau de pointe GÉANT qui dessert déjà 30 millions de chercheurs et d’étudiants, réalisent: collaborer, partager les ressources, accéder à l’information. Ces e-Infrastructures sont essentielles pour supporter l’innovation et les découvertes scientifiques nécessaires à l’UE dans son chemin pour le redressement économique.

C’est la raison pour laquelle, la Commission Européenne a adopté une Communication intitulée “Infrastructures TIC pour la science en ligne”. J’invite le parlement Européen, les 27 Etats Membres ainsi que les chercheurs et les industriels Européens concernés à en tirer parti pour stimuler leurs réflexions sur nos priorités futures.



**Viviane Reding**

Commissaire européen pour la société de  
l’information et des médias

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

**INFRASTRUCTURES TIC POUR LA SCIENCE EN LIGNE**

**1. INTRODUCTION**

**1.1. Objet de la présente communication**

La présente communication vise à *mettre en évidence* le rôle stratégique des infrastructures TIC<sup>1</sup>, atout majeur des politiques européennes en matière de recherche et d'innovation, et à *inviter* les États membres et les communautés scientifiques, en coopération avec la Commission européenne, à intensifier et conjuguer leurs efforts afin de promouvoir des infrastructures TIC de niveau international, également appelées *infrastructures électroniques*, ouvrant la voie aux découvertes scientifiques du 21<sup>e</sup> siècle.

**1.2. Contexte des infrastructures électroniques**

L'innovation, base du développement économique, est conditionnée par la rapidité des progrès scientifiques. Les activités scientifiques, de leur côté, reposent de plus en plus sur une collaboration ouverte, transnationale, entre chercheurs du monde entier. De plus, elles exigent de grands moyens de calcul de haute capacité permettant de modéliser des systèmes complexes et de traiter les résultats expérimentaux.

L'apparition de nouvelles méthodes de recherche tirant parti de moyens de calcul, de recueils de données et d'outils scientifiques avancés, en d'autres termes la *science en ligne*, est annonciatrice d'une révolution dans le processus de découverte scientifique, à l'instar de la «Renaissance scientifique»<sup>2</sup> qui a jeté les bases de la science moderne. Il est capital pour l'Europe d'épouser ce changement fondamental afin de conserver son avantage concurrentiel et de répondre aux besoins de la société.

Pour faciliter la transition rapide vers la science en ligne, la Commission européenne et les États membres ont réalisé des investissements importants dans les *infrastructures électroniques*, notamment dans le réseau paneuropéen de recherche GÉANT<sup>3</sup>, les réseaux scientifiques en grille, les infrastructures de données et le calcul intensif.

S'imposer comme le numéro un mondial de la science en ligne, faire des infrastructures électroniques un service d'utilité publique durable et les exploiter pour promouvoir l'innovation, tels sont les trois axes d'une stratégie européenne renouvelée pour soutenir les activités scientifiques novatrices à l'horizon 2020 et au-delà. Cette stratégie implique d'accomplir de grands progrès concernant le type et l'intensité des investissements, une plus grande interaction entre les politiques en matière de recherche et d'innovation et la coordination des stratégies nationales et communautaires.

---

<sup>1</sup> Technologies de l'information et des communications.

<sup>2</sup> M. B. Hall, *The scientific renaissance, 1450-1630*. ISBN 0486281159.

<sup>3</sup> Le réseau GÉANT fournit, en permanence, une gamme de services pas encore disponibles sur le marché (à un débit actuellement compris entre 40 et 100 gigabit/s) aux scientifiques de tous les pays.

### 1.3. Infrastructures électroniques et contexte politique

Le Conseil «Compétitivité»<sup>4</sup> a invité les États membres à *«encourager les établissements de recherche publics et privés à utiliser pleinement les systèmes distribués d'accès aux activités de recherche qui commencent à être mis en place («e-Science»), fondés sur des réseaux internationaux de recherche dont l'existence est rendue possible par la disponibilité et l'excellente qualité des infrastructures européennes de réseaux distribués, telles que les infrastructures GÉANT et les réseaux scientifiques en grille»,* d'où la nécessité accrue de coordonner les politiques.

Les infrastructures électroniques contribuent grandement à la réalisation des objectifs de la stratégie i2010<sup>5</sup> et au projet d'Espace européen de la recherche (EER<sup>6</sup>), et jouent un rôle essentiel de soutien au déploiement de nouvelles installations de recherche dont le développement s'articule autour des groupes stratégiques ESFRI<sup>7</sup> et e-IRG<sup>8</sup> dans le cadre d'un dialogue avec les États membres.

Le Conseil de Ljubljana<sup>9</sup> a remis l'accent sur le soutien en faveur de l'EER dont une nouvelle vision, a-t-il précisé, doit recouvrir la libre circulation de la connaissance («cinquième liberté») qu'il convient de faciliter en favorisant notamment l'accès à des infrastructures de recherche de classe mondiale ainsi que le partage et l'utilisation des connaissances par-delà les secteurs et les frontières. L'importance des infrastructures électroniques pour l'innovation est reconnue dans le rapport Aho<sup>10</sup> de mai 2008.

#### **i2010 (évaluation à mi-parcours, mai 2008)**

La contribution des TIC aux objectifs de Lisbonne est renforcée par le développement d'infrastructures électroniques (notamment GÉANT et réseaux scientifiques en grille) qui aident à mettre en place de nouveaux environnements de recherche, améliorant ainsi la productivité et la qualité des activités scientifiques.

#### **«Rapport Aho» (mai 2008)**

«Le développement réussi des [...] e-infrastructures a démontré l'importance de l'intervention européenne [...]. L'approche des «e-infrastructures» devrait s'étendre à des plateformes davantage axées sur les applications et sur les utilisateurs [...]. [Elles] doivent être mises en place dans des secteurs tels que l'administration en ligne (en particulier les marchés publics), la santé en ligne (applications transfrontières), la logistique et les transports.»

Le rapport met notamment en évidence la valeur ajoutée des infrastructures, de l'interopérabilité et des normes transnationales. L'étude ERINA<sup>11</sup> a confirmé le très fort potentiel qu'offrent les infrastructures électroniques, au-delà du domaine de la recherche, pour faciliter la mise sur le marché de technologies et services innovants.

<sup>4</sup> Conseil «Compétitivité» des 22-23 novembre 2007

([www.consilium.europa.eu/ueDocs/cms\\_Data/docs/pressdata/fr/intm/97288.pdf](http://www.consilium.europa.eu/ueDocs/cms_Data/docs/pressdata/fr/intm/97288.pdf)).

<sup>5</sup> Cadre politique de l'UE pour la société de l'information et les médias ([www.ec.europa.eu/i2010](http://www.ec.europa.eu/i2010)).

<sup>6</sup> COM(2007) 161: l'Espace européen de la recherche: nouvelles perspectives.

<sup>7</sup> European Strategy Forum on Research Infrastructures ([www.cordis.europa.eu/esfri](http://www.cordis.europa.eu/esfri)).

<sup>8</sup> e-Infrastructures Reflection Group ([www.e-irg.eu](http://www.e-irg.eu)).

<sup>9</sup> Conseil de Ljubljana de 2008 (<http://register.consilium.europa.eu/pdf/fr/08/st10/st10231.fr08.pdf>).

<sup>10</sup> Rapport Aho: «Information Society Research and Innovation: Delivering results with sustained impact», mai 2008

([http://ec.europa.eu/dgs/information\\_society/evaluation/rtd/fp6\\_ist\\_expost/index\\_en.htm](http://ec.europa.eu/dgs/information_society/evaluation/rtd/fp6_ist_expost/index_en.htm)).

<sup>11</sup> Étude ERINA ([www.erina-study.eu/homepage.asp](http://www.erina-study.eu/homepage.asp)).

La crise financière actuelle va mettre les budgets nationaux sous pression. Cependant, comme la Commission l'a récemment souligné<sup>12</sup>, il est désormais plus important que jamais de rechercher des modes de financement novateurs pour un large éventail de projets d'infrastructure, notamment dans les domaines des transports, de l'énergie et des réseaux de haute technologie.

## **2. LES INFRASTRUCTURES ELECTRONIQUES, ELEMENT DECLENCHEUR D'UNE NOUVELLE RENAISSANCE SCIENTIFIQUE**

### **2.1. Le changement fondamental de la science en ligne**

L'adoption des TIC pour toutes les phases du processus scientifique permettra aux chercheurs d'engager une collaboration rentable avec leurs homologues dans le monde, tandis que le recours accru à l'expérimentation *in silico*<sup>13</sup> ouvre de nouvelles perspectives à la coopération homme-machine et à la découverte scientifique. C'est ce que l'on appelle la transition des *laboratoires humides*<sup>14</sup> aux *environnements de recherche virtuels*, qui constitue l'aspect le plus évident du changement fondamental représenté par la science en ligne.

La systématisation de l'acquisition des connaissances par l'observation et l'expérimentation a été le facteur déterminant de la révolution scientifique à la Renaissance.

En portant l'expérimentation à des niveaux sans précédent pour explorer l'infiniment petit, l'infiniment grand et l'infiniment complexe, nous sommes à l'aube d'une nouvelle renaissance scientifique.

Par exemple, les recherches sur le changement climatique exigent des simulations informatiques complexes impliquant d'accéder à des données et de les récupérer dans des référentiels en ligne à travers le monde; le défi de créer des modèles individuels d'être humain pour des soins de santé personnalisés suppose des activités de modélisation et de simulation de plus en plus sophistiquées; reproduire le déroulement de phénomènes dangereux, comme les accidents nucléaires, les pandémies, les raz de marée, etc., obligent les chercheurs à réaliser de plus en plus leurs expériences dans des mondes

#### **Accélérer le processus de découverte de médicaments**

Lors de l'alerte à la grippe aviaire en 2006, les laboratoires européens et asiatiques ont utilisé 2 000 ordinateurs dans le réseau en grille EGEE<sup>15</sup> afin d'analyser 300 000 composants de médicament en 4 semaines – ce qui équivaut à 100 ans sur un seul ordinateur. Le criblage des médicaments *in silico* peut donc accélérer la découverte de molécules novatrices tout en limitant l'approche par essais et erreurs en laboratoire.

#### **Usines à données scientifiques**

Le grand accélérateur de hadrons, au CERN<sup>16</sup>, produit 600 millions de collisions de particules par seconde. Cela va générer d'énormes quantités de données qui seront mises à la disposition de 7 000 physiciens dans 33 pays par l'intermédiaire de GÉANT et d'infrastructures de science en ligne.

#### **Et si votre homologue était un robot?**

Les robots commencent à révolutionner les

<sup>12</sup> COM(2008) 800 final: un plan européen pour la relance économique.

<sup>13</sup> *In silico* est une expression utilisée pour désigner une expérience réalisée sur un ordinateur ou par simulation informatique, créée par analogie avec les expressions latines *in vivo* et *in vitro* qui sont utilisées pour désigner des expériences réalisées respectivement à l'intérieur et à l'extérieur d'organismes vivants.

<sup>14</sup> Un *laboratoire humide* est un laboratoire équipé des canalisations, de la ventilation et des équipements appropriés pour permettre les travaux pratiques de recherche scientifique.

<sup>15</sup> EGEE (Enabling Grids for E-science, [www.eu-egee.org](http://www.eu-egee.org)).

<sup>16</sup> CERN (Organisation européenne pour la recherche nucléaire).



virtuels plutôt que dans des environnements réels à haut risque où cela est coûteux.

La virtualisation des expériences permet aux chercheurs du monde entier de coopérer et de partager des données à l'aide de réseaux de recherche et d'infrastructures en grille avancés.

pratiques de laboratoire et à limiter la corvée des expériences manuelles en laboratoire humide. Ils permettent d'automatiser les processus et accélèrent la collecte et l'extraction de données scientifiques essentielles à la compréhension des phénomènes complexes et à la création de connaissances.

Ces changements ont pour effet de transformer les disciplines scientifiques en étendant leurs objectifs et leur champ d'application à d'autres domaines, et préfigurent une recherche interdisciplinaire.

Rester compétitif dans ce contexte de nouveaux défis scientifiques suppose la collaboration entre équipes de recherche en Europe et dans le monde, la capacité à utiliser et gérer des ensembles de données à la croissance exponentielle, et le recours à des environnements informatiques de haute performance pour la modélisation et la simulation.

Cela implique l'adoption généralisée de nouveaux environnements de recherche fonctionnant à l'aide de TIC avancées pour répondre efficacement aux besoins sans précédent des communautés scientifiques actuelles en matière de connectivité, de calcul et d'accès à l'information.

## **2.2. Les infrastructures électroniques pour la science en ligne d'aujourd'hui et de demain**

Les infrastructures électroniques, en facilitant les nouvelles découvertes scientifiques et l'innovation, constituent un moyen essentiel de soutenir la stratégie de Lisbonne pour une croissance durable et l'emploi.

Le programme-cadre de recherche et développement technologique (7PC) de la Commission européenne a grandement contribué au déploiement des infrastructures électroniques, non seulement pour accroître l'excellence scientifique, mais aussi pour promouvoir l'innovation et la compétitivité des entreprises.

Si l'Europe a affirmé sa suprématie dans le monde en ce qui concerne GÉANT et les réseaux scientifiques en grille, il reste encore à faire pour consolider sa position en matière de calcul intensif et pour élaborer une approche cohérente concernant l'accès aux données scientifiques et leur conservation.

La croissance exponentielle des performances du matériel informatique (doublement de la puissance de calcul tous les 18 mois, de la capacité de stockage tous les 12 mois et de la vitesse de connexion tous les 9 mois<sup>17</sup>) et les exigences scientifiques (atteindre l'échelle *exa*<sup>18</sup>) crée de nouveaux besoins et problèmes de conception des infrastructures électroniques de 2020.

### **Simulation en ingénierie à grande échelle**

La simulation informatique est essentielle à l'ingénierie moderne. La production d'artefacts élaborés comme les avions, les automobiles ou les appareils personnels repose sur des activités complexes de modélisation et de simulation et sur la coopération entre chercheurs et ingénieurs.

Les infrastructures électroniques doivent intégrer un ensemble plus étoffé de fonctionnalités, telles que systèmes et logiciels d'application de nouvelle génération, machines virtuelles, plateformes de fourniture de service, outils de visualisation, moteurs de recherche fondés sur la sémantique, etc., afin d'aider les équipes pluridisciplinaires à transformer les bits, octets et flops<sup>19</sup> en découvertes scientifiques et ingénierie complexe.

Il est à la fois nécessaire et possible de poursuivre le développement des infrastructures électroniques comme plateforme stratégique sur laquelle puisse se fonder la suprématie européenne dans le domaine de la science et de l'innovation. Cela demande que les États membres, la Commission européenne et les communautés scientifiques redoubtent d'efforts pour stimuler l'investissement dans les infrastructures électroniques et pour assurer la coordination et l'alignement requis des stratégies nationales et communautaire.

### **2.3. Une stratégie renouvelée**

La recherche en 2020 sans recours intensif aux infrastructures électroniques est unimaginable et l'Europe doit donc s'engager dans une stratégie renouvelée pour relever les défis et répondre aux priorités qui en découlent. Une telle stratégie repose sur trois grands axes interdépendants: la science en ligne, les infrastructures électroniques et l'innovation.

- Le premier axe suppose que l'Europe devienne un centre d'excellence pour la science en ligne, exploitant l'interdisciplinarité et la collaboration mondiale pour combiner les compétences et ressources complémentaires dans les simulations à forte intensité de calcul. L'Europe doit donc renforcer ses capacités de recherche en matière de calcul de haute performance.
- Le deuxième axe de la stratégie vise à consolider les infrastructures électroniques en tant que plateformes de recherche permanentes afin d'assurer la continuité de la recherche. La priorité est la fourniture, 24 heures sur 24 et 7 jours sur 7, de services de qualité production et la viabilité à long terme des infrastructures électroniques, ce qui exige de coordonner les efforts au niveau national et de l'UE et d'adopter des modèles de gestion appropriés.
- Le troisième axe concerne le potentiel d'innovation des infrastructures électroniques. Le transfert de compétences vers des domaines autres que la recherche (par exemple la santé en ligne, l'administration en ligne, l'apprentissage en ligne) et l'utilisation des infrastructures électroniques comme plateformes rentables pour des expériences techniques

<sup>17</sup> Lois communément admises régissant l'évolution de la technologie: Moore et Gilder.

<sup>18</sup> Des programmes de calcul à l'échelle *exa* (1 *exa* = 1 000 péta = 1 000 000 téra) à l'horizon 2020 font leur apparition au Japon et aux États-Unis.

<sup>19</sup> Les flops ou FLOPs – pour «FLoating point Operations Per second» ou opérations à virgule flottante par seconde – sont une unité de mesure de performance des ordinateurs.



à grande échelle (par exemple l'internet du futur, les logiciels massivement parallèles, les laboratoires vivants) sont autant de pistes à explorer.

Cette stratégie sera mise en application par une série de mesures concrètes axées sur les différents domaines structurels des infrastructures électroniques. Le succès de sa mise en œuvre dépend de la coordination des efforts et d'un engagement accru de la part des autorités de financement nationales et de l'UE.

### 3. L'EUROPE OUVRE LA VOIE

#### 3.1. Les infrastructures électroniques aujourd'hui

Les infrastructures électroniques sont aujourd'hui organisées en cinq domaines étroitement liés et assurent ensemble une grande diversité de fonctions et de services.

- **GÉANT** est le plus grand réseau mondial de communication multi-gigabit destiné à la recherche et l'enseignement. En Europe, GÉANT est déjà utilisé par près de 4 000 universités et centres de recherche, et relie 34 réseaux nationaux pour la recherche et l'enseignement (NREN). Il est connecté à des réseaux similaires dans le monde (régions des Balkans, de la mer Noire et de la Méditerranée, Asie, sud de l'Afrique et Amérique latine) formant ainsi un unique réseau mondial de recherche. La prééminence de GÉANT résulte d'un modèle de gestion consolidé selon lequel les NREN assurent le déploiement requis au niveau national et coordonnent collectivement la mise en œuvre du réseau paneuropéen par l'alignement des options stratégiques et technologiques ainsi que par la mise en commun des ressources financières au niveau national et européen.

#### Qu'est-ce qu'une infrastructure électronique?

Une infrastructure électronique est un environnement dans lequel les moyens de recherche (matériel, logiciels et contenu) peuvent être aisément partagés et utilisés chaque fois que c'est nécessaire pour obtenir des résultats plus probants.

Un tel environnement comprend des réseaux, réseaux en grille et intergiciels, des moyens de calcul, des bancs d'essai expérimentaux, des référentiels de données, des outils et instruments, et le soutien opérationnel pour une collaboration scientifique virtuelle mondiale.

#### Qu'est-ce qu'un réseau en grille?

Un réseau en grille est un service de partage de la puissance de calcul et de la capacité de stockage des données sur internet. Cela dépasse largement les connexions entre ordinateurs et vise, en dernier lieu, à transformer le réseau mondial d'ordinateurs en une immense ressource de calcul permettant le fonctionnement, à grande échelle, d'applications à forte intensité informatique et de données.

- Les **réseaux scientifiques en grille** ont fait leur apparition pour répondre aux besoins des disciplines scientifiques les plus exigeantes (par exemple la physique des hautes énergies, la bio-informatique), partager et combiner la puissance des ordinateurs et d'instruments scientifiques sophistiqués souvent uniques. Grâce au soutien des programmes-cadres de l'UE, l'Europe héberge aujourd'hui les plus grands réseaux pluriscientifiques en grille. L'EGEE exploite à ce jour un réseau pluridisciplinaire en grille de plus de 80 000 ordinateurs sur 300 sites dans 50 pays à travers le monde, au service de plusieurs

milliers de chercheurs. Le projet DEISA<sup>20</sup>, par l'interconnexion des 11 supercalculateurs les plus puissants du continent, offre un environnement de calcul intensif permanent et de qualité production en Europe.

- Dans le domaine des **données scientifiques**, le but est d'enrayer la prolifération accélérée et incontrôlée des données qui, si elle n'était pas maîtrisée, pourrait nuire à l'efficacité du processus de découverte scientifique<sup>21</sup>. Il est donc capital de mettre au point de nouveaux outils et méthodes pour assurer la disponibilité, le traitement et la conservation de quantités de données sans précédent. Concernant les référentiels de données, la situation est assez variable en Europe, mais il existe une base solide pour l'élaboration d'une stratégie cohérente afin de remédier au morcellement et de permettre aux communautés scientifiques de mieux gérer, utiliser, partager et conserver les données. Les projets financés par l'Europe dans le domaine des infrastructures de données scientifiques partagent une même vision: toute source de contenu scientifique (rapports scientifiques, articles de chercheurs, données expérimentales ou d'observation, médias enrichis, etc.) doit être aisément accessible en tant que plateforme de partage des connaissances au moyen de services conviviaux d'infrastructure électronique.
- Les infrastructures électroniques de **calcul intensif** répondent aux problèmes, complexes et à forte intensité de données, posés par la fourniture des nouveaux moyens de calcul et de simulation dont la science moderne a besoin. L'intérêt stratégique des États membres et de la communauté scientifique pour des services européens de calcul et de simulation de haute performance a conduit à la création d'une nouvelle infrastructure électronique, PRACE<sup>22</sup>, bénéficiant d'un soutien au titre du programme Capacités du septième programme-cadre de recherche.
- Les **communautés scientifiques virtuelles mondiales**, anticipant l'avènement des modèles de recherche 2.0<sup>23</sup>, ont ouvert de nouvelles perspectives de collaboration pluridisciplinaire transnationale entre communautés scientifiques. Un changement culturel est en train de se produire dans la façon dont les connaissances scientifiques sont produites et diffusées, qui provoque l'apparition de communautés scientifiques virtuelles mondiales. L'Europe contribue déjà à l'innovation dans le processus scientifique en permettant aux communautés scientifiques d'utiliser des infrastructures électroniques pour relever les grands défis de la recherche au niveau mondial.

### 3.2. Les infrastructures électroniques pour 2020 et au-delà

La réponse de l'Europe aux problèmes à long terme que pose la science en ligne exige une approche plus efficace et plus coordonnée des investissements européens dans les infrastructures scientifiques de niveau international. En apportant des réponses communes aux différents besoins des utilisateurs, les infrastructures électroniques s'avèrent capitales pour promouvoir l'excellence scientifique, encourager les partenariats scientifiques mondiaux et favoriser le développement d'un capital humain de haute qualité, tout en permettant des économies d'échelle. Les infrastructures électroniques sont des biens publics qui étayent les politiques en matière d'enseignement, de recherche et d'innovation. La participation active des pouvoirs publics à la définition des priorités et stratégies est donc essentielle.

---

<sup>20</sup> DEISA (Distributed European Infrastructure for Supercomputing Applications, [www.deisa.eu](http://www.deisa.eu)).

<sup>21</sup> COM(2007) 56: l'information scientifique à l'ère numérique.

<sup>22</sup> PRACE (Partnership for Advanced Computing in Europe, [www.prace-project.eu](http://www.prace-project.eu)).

<sup>23</sup> La recherche 2.0 est un terme désignant l'utilisation de la technologie du web 2.0 pour développer la créativité, le partage d'informations et la collaboration dans le domaine de la recherche.

La capacité unique de **GÉANT** de permettre une collaboration scientifique novatrice grâce à une connectivité à haut débit et des services avancés constitue l'une des réussites européennes les plus marquantes. Pour que l'Europe défende sa grande tradition d'innovation et de découverte scientifique au-delà de 2020, GÉANT doit mettre à profit ses extraordinaires performances pour atteindre l'échelle *exa* et contribuer à la conception de l'internet du futur.

Aujourd'hui, la viabilité des **réseaux scientifiques en grille** dépend surtout de la forte demande des communautés d'utilisateurs scientifiques collaborant à des projets financés au titre de programmes nationaux et communautaires. Cela présente un risque d'interruption opérationnelle et finit par devenir un frein à la pleine exploitation des réseaux en grille.

#### **Initiatives nationales de réseau en grille (NGI)**

Les NGI sont des entités investies d'une mission de service public visant à intégrer des sources de financement au niveau national pour la fourniture de services fondés sur les réseaux en grille. Elles fournissent plusieurs de ces services à des communautés scientifiques nationales par l'intermédiaire d'un guichet unique.

Les cycles de développement technologique courts, basés sur des projets, peuvent nuire à l'interopérabilité des infrastructures en grille et donc empêcher la coopération interdisciplinaire et les économies d'échelle. Les projets EGEE et DEISA ont déjà permis de beaucoup progresser dans l'association des disciplines et la coordination des stratégies. Dans un souci de viabilité à long terme, ces efforts doivent prendre la forme de modèles d'organisation réellement paneuropéens qui ouvriront les infrastructures électroniques en grille à toutes les disciplines scientifiques et compléteront les stratégies nationales de financement en faveur de la science en ligne. Plusieurs **initiatives nationales de réseau en grille** font leur apparition pour répondre de façon coordonnée et rentable aux besoins des disciplines scientifiques en matière de moyens de calcul.

L'objectif des **infrastructures électroniques de données scientifiques** est de créer un environnement de référentiels numériques européens en combinant les référentiels nationaux et disciplinaires et en y apportant une valeur ajoutée pour répondre aux demandes d'amélioration de l'accès à l'information scientifique émanant des États membres.

#### **Des données, des données et encore des données...**

La taille des référentiels de bio-informatique augmente à une vitesse exponentielle. D'ici à 2012, les informations qui sont accumulées chaque année dans un seul référentiel de données représenteront 4 pétaoctet/an, soit une pile de CD de 10 km de haut.

L'apparition d'une science des gros volumes de données est un phénomène mondial<sup>24</sup> reflétant l'importance croissante des données d'observation et expérimentales brutes dans pratiquement tous les domaines scientifiques (sciences humaines, biodiversité, physique des hautes énergies, astronomie, etc.). L'Europe doit accorder une attention particulière à l'accessibilité, l'assurance qualité et la conservation des principales collections de données. Par exemple, les politiques européennes en matière d'environnement reposent sur la directive INSPIRE<sup>25</sup> qui vise à mettre en place une infrastructure européenne d'information géographique afin de fournir des services intégrés de données géographiques. Le contexte en matière de données numériques étant disparate – on estime que seulement 28% des résultats de la recherche sont gérés dans

<sup>24</sup> Programme DataNet de la Fondation nationale pour la science des États-Unis (<http://www.nsf.gov/pubs/2008/nsf08021/nsf08021.jsp>).

<sup>25</sup> Directive 2007/2/CE: infrastructure d'information géographique dans la Communauté européenne.

des référentiels numériques<sup>26</sup> –, il faut élaborer une nouvelle stratégie de gestion de l'information scientifique et des politiques associées reposant sur les activités d'exploration des principaux acteurs de la recherche (par exemple LEBM, ASE, CEPMMT, CERN<sup>27</sup>) ainsi que des établissements et bibliothèques universitaires.

Le **calcul intensif** a été défini comme une priorité absolue pour accroître les performances scientifiques de l'Europe. Cela exige une nouvelle stratégie visant à la participation des entreprises et à la coordination entre autorités de financement<sup>28</sup>. Couvrant les problèmes stratégiques, politiques, techniques, financiers et de gestion liés au calcul intensif, PRACE contribue largement à mobiliser des fonds nationaux importants en vue de créer un environnement de machines à l'échelle *péta* en Europe, avec l'exaflops en ligne de mire d'ici à 2020.

Pour soutenir efficacement la science en ligne et prendre la tête des **communautés scientifiques virtuelles mondiales**, l'Europe doit continuer à développer des infrastructures de niveau international capables de prendre en charge de nouveaux modèles participatifs. Cela offre une occasion unique de renforcer le rôle de la recherche européenne dans un contexte mondial en pleine évolution.

Toutefois, pour exploiter pleinement le potentiel de la collaboration scientifique mondiale, il est nécessaire de régler plusieurs problèmes. Il s'agit du choc culturel entre les différentes disciplines, de la nécessité de repenser les modèles d'organisation, et de l'instauration de mécanismes d'assurance qualité et de modèles économiques.

Il est également indispensable d'élaborer de nouvelles stratégies de développement technologique des infrastructures électroniques pour obtenir des solutions à l'épreuve du temps, fondées sur des normes ouvertes, qui puissent être maintenues et perfectionnées à long terme et apporter une valeur ajoutée aux investissements dans les installations de recherche, les instruments de grande envergure ou uniques, etc.

#### 4. ACTIONS DE L'EUROPE

La réussite d'une stratégie renouvelée dépend de l'application d'une série de mesures concrètes ciblant différents aspects des infrastructures électroniques européennes, et de la création de synergies entre ces mesures.

##### 4.1. Réaffirmer la suprématie de GÉANT à l'échelle mondiale

GÉANT doit, en étroite concertation avec les NREN, continuer à fournir une connectivité haut de gamme permanente, avec des niveaux de performance très supérieurs, aux chercheurs, enseignants et étudiants afin de faciliter l'accès aux ressources et équipements distribués. Il doit accroître sa dimension mondiale en couvrant à la fois des régions développées et en développement<sup>29</sup>.

---

<sup>26</sup> *Investigative Study of Standards for Digital Repositories and Related Services DRIVER* (<http://dare.uva.nl/document/93727>).

<sup>27</sup> LEBM (Laboratoire européen de biologie moléculaire), ASE (Agence spatiale européenne), CEPMMT (Centre européen pour les prévisions météorologiques à moyen terme), CERN (Organisation européenne pour la recherche nucléaire).

<sup>28</sup> L'Europe était sous-représentée dans les classements mondiaux illustrant l'évolution du calcul de haute performance (<http://www.top500.org/>).

<sup>29</sup> En s'inspirant d'initiatives comme ALICE (<http://alice.dante.net>), EUMEDconnect ([www.eumedconnect.net](http://www.eumedconnect.net)), TEIN2 ([www.tein2.net](http://www.tein2.net)) soutenues par les DG RELEX, DEV et AIDCO.

GÉANT doit également intégrer les dernières évolutions technologiques en matière de travail en réseau et prendre en charge l'expérimentation de nouveaux modèles préfigurant l'internet du futur<sup>30</sup>.

**Les États membres sont invités à renforcer la coordination des politiques nationales et européennes dans le domaine des réseaux pour la recherche et l'enseignement.**

**Les États membres et les communautés scientifiques sont invités à soutenir et utiliser GÉANT comme plateforme expérimentale préfigurant l'internet du futur.**

**La Commission, par le 7PC et la coopération internationale, continuera à apporter un soutien résolu à GÉANT pour accroître ses capacités et sa dimension mondiale.**

#### **4.2. Structurer les réseaux scientifiques en grille**

À l'avenir, les réseaux scientifiques en grille européens devront tirer parti du succès des initiatives actuelles, justifiées par les besoins communs de différentes disciplines scientifiques, et faire en sorte que les entreprises les adoptent.

Cependant, pour une plus grande viabilité à long terme, les modèles de gestion doivent évoluer dans le sens d'une initiative européenne de réseau en grille (EGI) s'inspirant des initiatives nationales de réseau en grille (NGI) naissantes.

**Les États membres sont invités à développer et poursuivre les initiatives nationales de réseau en grille (NGI) en tant que base d'une stratégie européenne renouvelée.**

**La Commission favorisera la transition vers de nouveaux modèles de gestion des réseaux scientifiques en grille européens ainsi que leur déploiement effectif au service d'un large éventail de domaines scientifiques, en assurant l'interopérabilité technologique avec les réseaux en grille mondiaux.**

#### **4.3. Améliorer l'accès à l'information scientifique**

Les infrastructures électroniques européennes et nationales doivent relever le nouveau défi d'une science centrée sur les données. Pour y parvenir, l'Europe doit créer un environnement, géré de façon cohérente, de référentiels d'information scientifique. L'Europe doit élaborer des politiques cohérentes pour améliorer l'accès à l'information scientifique (par exemple selon les indications du document exposant la position de l'ESFRI sur les données scientifiques, la communication sur «l'information scientifique à l'ère numérique: accès, diffusion et préservation»<sup>31</sup> et le projet pilote de libre accès, dans le cadre du 7PC<sup>32</sup>, lancé en 2008).

**Les États membres et les communautés scientifiques sont invités à accroître les investissements dans les infrastructures de données scientifiques et à promouvoir le partage des meilleures pratiques.**

**La Commission intensifiera, au titre du 7PC, les investissements déclencheurs dans les infrastructures de données scientifiques pour soutenir les politiques en matière d'accessibilité et de conservation.**

---

<sup>30</sup> En soutenant des initiatives comme FIRE (Future Internet Research & Experimentation, <http://cordis.europa.eu/fp7/ict/fire/>).

<sup>31</sup> COM(2007) 56: communication sur l'information scientifique à l'ère numérique: accès, diffusion et préservation.

<sup>32</sup> <http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1680>.

## **5. CONCLUSIONS**

Il est capital de soutenir les politiques en matière de recherche et d'innovation pour que l'Europe puisse relever les formidables défis des 10 à 15 années à venir. En effet, l'activité scientifique va connaître de grands changements dans la façon dont elle s'exerce. En s'attaquant à des problèmes scientifiques ayant un impact sociétal mondial, les chercheurs seront confrontés à des degrés de complexité sans précédent et il sera essentiel de regrouper les connaissances tirées des différents domaines de la recherche.

Les infrastructures électroniques fournissent les plateformes de base aux applications à forte intensité de calcul qui permettent une collaboration combinant les connaissances de diverses disciplines scientifiques. De l'utilisation d'environnements de réseau hautement distribués, comme GÉANT, naîtront de nouvelles formes d'organisation, y compris virtuelles, à travers le globe.

Les efforts accrus et conjugués des États membres, de la Commission européenne et des communautés scientifiques concernées accéléreront le déploiement des infrastructures électroniques afin d'accroître leurs capacités et fonctionnalités de plusieurs ordres de grandeur.

La stratégie renouvelée pour s'imposer comme le numéro un de la science en ligne, mettre au point des infrastructures électroniques de niveau international et exploiter le potentiel d'innovation de la recherche, est essentielle pour faire de l'Europe un centre d'excellence scientifique et un partenaire de dimension réellement mondiale dans ce domaine.



# EXAMPLES OF e-INFRASTRUCTURES PROJECTS<sup>1</sup>

## • Linking the ideas at the speed of light: GÉANT

GÉANT2, the world's largest multi-gigabit communication network dedicated to research and education, offers the scientists the possibility to store, transmit and analyse more data than ever before.



GN2 ( [www.geant2.net](http://www.geant2.net) )



With the EU's commitment to global outreach and vision in creating a single global research network for international collaboration, GN2, the e-Infrastructure project responsible for operating GÉANT2, the high bandwidth, pan-European network, has played a key role in shaping global research and education over the past four years. The GN2 project delivers high bandwidth network performance through the world-leading GEANT2 network and a range of associated multi-domain services to optimise the user experience.

By connecting researchers across Europe, through the national research and education networks (NRENs), GÉANT2 enables the best minds across Europe and the world to together pursue collaborative and ground-breaking research activities that were previously impossible or inefficient due to limited and poor quality data connections.

Together with the NRENs it connects, GÉANT2 provides the fastest connection speeds and widest geographical reach of any network of its kind in the world, with outstanding levels of reliability. GÉANT2 has also extended its reach and has narrowed the global digital divide with its links to other EU-funded initiatives in developing world regions. As a world-leading research network, years ahead of those provided by commercial counterparts, GÉANT2 provides enhanced bandwidth, point-to-point links and a range of multi-domain services to optimise the user experience and to help keep researchers at the forefront of their respective fields of research, retaining and developing academic expertise in Europe and ultimately boosting Europe's competitiveness.

GÉANT2 replaced GÉANT in mid 2004 and GÉANT3 will be launched towards the end of 2009. It allows more data to be stored, transmitted and analysed than ever before.

<sup>1</sup> Cette section de la brochure est disponible en anglais uniquement.

- **Sharing the best scientific resources: e-Science grids**

Grid infrastructures make unprecedented amount of computing power available to scientists by pooling distributed resources across Europe and beyond in a secure and reliable way. Using grids helps researchers to solve complex scientific problems more effectively and efficiently than it was possible before.



**EGEE-III** (Enabling Grids for E-scienceE, [www.eu-egee.org](http://www.eu-egee.org))



EGEE, the world's largest multidisciplinary grid infrastructure for research project, connects approximately 300 sites in 50 countries to produce a reliable and scalable computing resource for the global science community. By giving researchers in academia and business access to a production level grid infrastructure, with 100.000 processing cores and secure data transfer and storage, independent of their geographic location, EGEE enables research that would not otherwise be possible.

EGEE supports applications from many scientific domains, such as astronomy, astrophysics, life sciences, computational chemistry, earth sciences, high energy physics, finance and fusion. In addition, there are several applications from business sectors running on the EGEE Grid, such as geophysics and the plastics industry. The project develops and deploys the open source, business-friendly middleware distribution "gLite" that virtualises separate grid resources and allows services like data management, complex workflow tools and discipline specific components to run on top of them.

EGEE operates with two clear goals fundamental to European research infrastructures: (1) expanding and optimising its services for researchers who already depend upon the grid for their daily work, while reaching new user communities, and (2) working towards a sustainable, federated European Grid Initiative. By supporting the long-term availability of grid computing, EGEE project strengthens the ability of ICT infrastructures to transform the way science is carried out and thereby increases the impact of the European Research Area.



DEISA pulls together the power of the 12 largest supercomputers in Europe, providing virtual access to these large high-performance computing facilities to solve scientific problems of huge complexity. The DEISA Consortium has deployed and operated the Distributed European Infrastructure for Supercomputing Applications. It continues to support and further develop the distributed high performance computing (HPC) infrastructure and its services through the DEISA-2 project funded until April 2011. All HPC relevant activities and services such as Application Enabling, Operation, and Technologies are continued and further enhanced.

The DEISA Extreme Computing Initiative (DECI) regularly attracts the most challenging computational science projects Europe-wide. During the last four years, DECI has supported scientists from over 160 universities or research institutes from 15 European countries and four other continents, as such the continuation of DECI is crucial. To further enhance the impact of the European supercomputer infrastructure, the DEISA consortium has extended the single project support to also supporting European science communities globally. Furthermore collaborative activities are carried with other computing infrastructure initiatives world-wide.

The cooperation with the PRACE project which is preparing for the installation of a limited number of leadership-class supercomputers in Europe is of strategic importance.

## • Accessing knowledge: Scientific data

By weaving together scientific data repositories, multidisciplinary research communities are given ready access to an unprecedented amount of data. This allows for greater creativity and productivity in research and a broader dissemination of information.



**DRIVER-II** (Digital Repository Infrastructure Vision for European Research, [www.driver-repository.eu](http://www.driver-repository.eu))



Considered the largest initiative of its kind in helping to enhance repository development worldwide, DRIVER has a clear vision: “All research institutions in Europe and world wide make all their research publications openly accessible”. This follows the Berlin Declaration, which called in 2003 for “free and unrestricted access to sciences and human knowledge representation worldwide”.

DRIVER is a multi-phase effort whose vision and primary objective is to create a cohesive, robust and flexible, pan-European infrastructure for digital repositories, offering sophisticated services and functionalities for researchers, administrators and the general public.

DRIVER has established a network of relevant experts and Open Access repositories with a fully functional, state-of-the art service, extending the network to a larger confederation of repositories.

DRIVER II is to expand, enrich, and strengthen the results of the initial DRIVER project by providing a production-quality digital repository infrastructure and establishing a European digital repositories confederation. The DRIVER Infrastructure acts as a showcase for repository developments, as a networking tool for the Driver Community and demonstrates a range of end-user services. The DRIVER Confederation comprises data and service providers, the major repository software platforms, representatives of emerging standards, protocols and data models, and NGO's serving the repository community.



The wide range of scientists trying to mitigate climate change and its impact on society need to access climate data, in particular data from model projections for the next decades and centuries, and to assess the suitability of that data for their purposes. Currently, even if large repositories of climate data are accessible, these repositories are poorly connected and described. The data is therefore difficult to understand, compare and exploit because there is no common information paradigm.

METAFOR's main objective is to define a Common Information Model (CIM) that describes in a standard way climate data and the models that produce the data. METAFOR builds on existing metadata (data describing data) currently used in existing data repositories and addresses issues like metadata fragmentation, gaps, or duplication. In close interaction with related initiatives at the international level, METAFOR is proposing solutions to identify, access and use the climate data in these repositories. METAFOR is optimising the way climate data infrastructures are used to store knowledge, thereby adding value to primary research data for an increasingly wide range of stakeholders.

METAFOR is taking the first step in doing for climate data what search engines have done for the Internet: it is putting users of climate data in touch with the information they need. Simply unlocking existing climate data repositories for new communities will provide new exciting opportunities for scientific research, policy making, and private sector competitiveness.

- **Designing future facilities: Novel e-infrastructures**

**Design Studies:** European scientific communities need new e-Infrastructures to remain at the forefront of the global research. The studies for design and creation of new infrastructures of pan-European dimension will pave the way to unique services to scientists, thus enlarging the communities and structuring the European Research Area.



**EGI\_DS** (The European Grid Initiative (EGI) Design Study, <http://web.eu-egi.eu>)



The European Grid Initiative (EGI) aims to develop a sustainable pan-European grid infrastructure, for optimal sharing of computing and data resources.

The EGI is a partnership between National Grid Initiatives (NGIs) and a coordinating body, the EGI Organization (EGI.org), and will begin full operation in 2010. It stems from a series of recent grid projects, such as the EGEE projects, which have contributed to the current leading position of Europe in e-Infrastructures, through the development of a large-scale production-quality grid for scientists.

The EGI goes a step further by moving from a structure based upon short-term funded projects to a long-term service, with the creation of a new organizational model that will coordinate the existing resources and ensure the development of a sustainable grid infrastructure, enabling Europe to secure its leadership in e-Infrastructures.

The EGI\_DS aims to define and implement the new organization's activities, and coordinates the efforts of the participating countries as they move towards the EGI model.

EGI\_DS has already published a series of Deliverables describing the individual components of the EGI infrastructure. The full organization's operational and financial framework is defined in the EGI Blueprint, endorsed by the NGI representatives on January 20th, 2009.



***Construction of new infrastructures:*** The aim is to create a new research infrastructure in the area of High Performance Computing (HPC) in order to remain at the forefront of scientific research and to help industry to strengthen its technological know-how. This action implements one of the projects foreseen in the European Strategy Forum on Research Infrastructures (ESFRI) roadmap.

**PRACE** (Partnership for Advanced Computing in Europe, [www.prace-project.eu](http://www.prace-project.eu))



PRACE is preparing the creation of a persistent world-leading high-end computing infrastructure for Europe.

European scientists and engineers will use the services to attain global leadership in research and development and to advance Europe's competitiveness.

Access will be granted through a peer review process. Several world-class centres will provide a variety of architectures to meet the needs of the different scientific and industrial domains and applications.

The PRACE project undertakes all legal, administrative and technical work to establish the infrastructure as a single legal entity starting in 2010. Prototypes of potential production systems have been selected based on an in-depth analysis of user requirements and existing mature technological solutions and are being evaluated. A European benchmark suite has been created for the selection of the first petascale production computer systems and researchers are being trained in the use of the future systems.

- **Innovating the scientific process: global virtual research communities**

By their ability to solve complex problems, remove geographic distance and make optimal use of expensive instruments, ICT based e-Infrastructures are changing the way research communities address key societal and economic issues, such as energy, healthcare or climate change.



**GENESI-DR** (Ground European Network for Earth Science Interoperations-Digital Repositories, <http://www.genesi-dr.eu>)



Earth scientists need access to loads of data in the shortest possible time to perform their research and produce trustworthy results based on all relevant sources of available information. The discovery of these data may be not trivial as data can be distributed across differently organised archives. Accessing them is even more difficult since different archives may use different interfaces. When these problems are solved, earth scientists should know how to interpret these data and have availability of applications, tools and services for their processing and analysis.

GENESI-DR aims to overcome these problems. It provides a framework where data repositories can be easily integrated and users are given a single access point for data discovery and access. GENESI-DR offers as well processing power, and the possibility to incorporate processing services so supporting the users in the analysis of the data.

The GENESI-DR e-Infrastructure is currently being validated against specific applications in the land, atmosphere and marine domain, providing initial processing services to expert users. GENESI-DR is also analysing common approaches to the preservation of data for an effective future access of the Earth Science community knowledge.



EXPReS makes e-VLBI available to astronomers by connecting radio telescopes around the world to high-speed fiber optic networks.

In e-VLBI, telescopes simultaneously observe a cosmic source. The data is streamed in real-time to a supercomputer which processes the data used to create the highest resolution images of astronomical objects. Historically, VLBI data was stored on hard disks and mailed to a computing facility, requiring months of waiting before results were available.

e-VLBI's real-time capability allows astronomers to react quickly to transient events by conducting nearly immediate follow-on observations.

As of March 2009, EXPReS has made e-VLBI with up to nine simultaneously operated and connected telescopes available to astronomers.

---

Commission européenne

INFRASTRUCTURES TIC POUR LA SCIENCE EN LIGNE

Luxembourg: Office des publications de l'Union européenne

2009 — 24 pp. — 21 x 29,7 cm

ISBN 978-92-79-11285-0

doi 10.2759/42447

# INFRASTRUCTURES TIC POUR LA SCIENCE EN LIGNE

## Commission européenne

Direction générale Société de l'information et médias

Unité GÉANT & infrastructure électronique

[http://cordis.europa.eu/fp7/ict/e-infrastructure/home\\_en.html](http://cordis.europa.eu/fp7/ict/e-infrastructure/home_en.html)



Office des publications

doi 10.2759/42447

ISBN 978-92-79-11285-0



9 789279 112850

