

# Plateformes d'envergure en Sciences de l'Information

Conseil Scientifique d'Institut de l'INS2I

Fiche de synthèse – 24 novembre 2015, v1

Rédigée par le groupe de travail Plateformes :

*Michel Beaudouin-Lafon (animateur), François Chaumette, Julien Gossa, Inbar Fijalkow, Anne-Catherine Letournel, Serge Torres, Hélène Touzet*

## 1. Introduction

Les recherches en sciences de l'information nécessitent de plus en plus l'utilisation de plateformes matérielles et logicielles spécialisées. Cela est dû à plusieurs facteurs, parmi lesquels :

- Une approche expérimentale de plus en plus souvent nécessaire pour valider des résultats de recherche (analyse de performance par exemple), mais aussi pour conduire l'ensemble du processus de recherche (interaction humain-machine par exemple) ;
- Le passage à l'échelle imposé par l'évolution des enjeux scientifiques dans de nombreux domaines, qui nécessite des infrastructures dédiées (HPC, big data, etc.) ;
- L'importance croissante des recherches qui combinent le monde physique et le monde numérique (robotique, internet des objets, etc.).

En conséquence, un nombre conséquent de *plateformes d'envergure*, c'est-à-dire qui dépassent le cadre d'une équipe ou d'un laboratoire, ont vu le jour depuis quelques années dans les unités de l'Institut, dont certaines dans le cadre des appels EQUIPEX du PIA.

Le CSI de l'INS2I a conduit le 1<sup>er</sup> juin 2015 une journée d'auditions et de réflexion sur le thème des plateformes. Afin de focaliser la réflexion, cette journée n'a pas porté sur les plateformes purement logicielles, et n'a que peu abordé les plateformes de calcul ou de stockage, dont l'importance pour la recherche en informatique est désormais bien établie. Les présentations et les réflexions ont donc porté surtout sur les plateformes ayant une forte composante matérielle et d'envergure régionale ou nationale, voire internationale. Ont été auditionnés :

- Jean-François Gibrat, directeur de l'IFB, pour l'infrastructure nationale de bioinformatique
- Eric Fleury, professeur à l'ENS Lyon, pour la plateforme FIT (Internet des objets)
- Michel de Mathelin, directeur de iCube, pour l'Equipex ROBOTEX
- Jean-Marc Vézien, ingénieur au LIMSI, pour l'Equipex DIGISCOPE
- Vivien Quema, professeur au LIG, pour la plateforme d'usages DOMUS

## 2. Bilan des réflexions

Les auditions ont montré la diversité des plateformes, qu'il s'agisse des types d'équipements déployés, de leur portée géographique, des publics visés, ou des moyens, notamment humains, de support. Elles ont montré également un certain nombre de points communs.

### 2.1 Finalités des plateformes

D'une part ces plateformes sont très souvent utilisées à la fois comme *outils* de recherche et comme *objets* de recherche. En tant qu'outil, une plateforme fournit une infrastructure permettant aux chercheurs de réaliser leurs travaux, au même titre qu'un

microscope pour un biologiste. En tant qu'objet de recherche, une plateforme est conçue et développée dans le cadre d'un projet de recherche pour permettre certains types de travaux et apporter des réponses à des questions de recherche, mais aussi pour poser de nouvelles questions, selon une démarche exploratoire. Ces deux finalités ne sont pas exclusives, et peuvent évoluer au cours de la vie d'une plateforme : un outil peut devenir un objet de recherche en l'adaptant ou en le modifiant, et un objet de recherche peut se stabiliser et, même s'il devient obsolète pour cette finalité, devenir un outil utile pour d'autres chercheurs. Il est donc important de considérer une plateforme non pas seulement comme une ressource matérielle avec des objectifs et des utilisateurs bien identifiés, mais comme une construction socio-technique de ses différents acteurs amenée à évoluer dans le temps (certains diraient un « écosystème »).

Dans certains domaines, les plateformes peuvent ainsi contribuer significativement à l'animation et la structuration des communautés en favorisant les rencontres et les échanges autour d'outils partagés. Ce facteur va de pair avec un changement dans les rapports qu'entretiennent les plateformes et leurs usagers. On passe d'une relation client-fournisseur à une co-construction entre les gestionnaires et les utilisateurs, d'une définition statique à une vision dynamique de l'organisation et des prestations offertes.

Par ailleurs, un certain nombre de plateformes se trouvent à la limite interdisciplinaire entre les sciences de l'information et d'autres disciplines. La question se pose alors de l'intérêt d'un tel équipement pour la *recherche* en sciences de l'information. L'intérêt peut venir de la possibilité d'observer l'utilisation réelle d'une technologie et des données ainsi recueillies, qui peuvent alimenter des recherches futures. Il peut aussi venir d'un travail de co-conception de la plateforme, bénéfique aux différents partenaires. Il semble que nous soyons ici dans les questionnements traditionnels de l'interdisciplinarité, qui ne peuvent être tranchés que par une analyse fine de chaque situation.

D'un autre côté, les plateformes en sciences de l'information peuvent jouer un rôle de ressource et d'expertise au service de la construction des outils numériques dont se dotent les différents secteurs de la communauté scientifique—voir par exemple le rôle de Grid 5000 dans le développement de l'infrastructure pour la bioinformatique. Il s'agit alors d'une forme de valorisation.

Enfin, les plateformes sont également des vecteurs de diffusion de la connaissance et des supports à la formation. En informatique, où l'invisibilité du logiciel et des algorithmes qui le sous-tendent rend souvent difficile la communication sur nos pratiques et résultats de recherche, une plateforme matérielle fournit un point d'ancrage commode pour illustrer une problématique ou un résultat. Mais cela conduit aussi parfois les plateformes à devenir des « usines à démos », avec un coût (en temps) conséquent pour les chercheurs. Pour la formation, les plateformes peuvent être un élément d'attractivité pour les étudiants et, là aussi, un élément tangible des savoirs et des connaissances transmises. Mais le partage de ces ressources entre recherche et enseignement peut aussi être source de difficultés.

## *2.2 Organisation et gestion des plateformes*

Les plateformes considérées sont très souvent organisées en réseau et sont souvent ouvertes à des utilisateurs au-delà des partenaires de la plateforme. L'organisation en réseau, à une échelle géographique plus ou moins grande, apparaît comme une nécessité afin de justifier les investissements par un plus grand nombre d'utilisateurs, de mutualiser les moyens, et d'être compétitif au niveau international. L'organisation en réseau rend aussi une plateforme plus robuste. L'ouverture répond à la fois à un souci de visibilité et de partage, et parfois à une demande des bailleurs de fonds. Elle est aussi

un moyen de développer les transferts technologiques, en donnant accès à des démonstrateurs technologiques à des acteurs industriels. Cependant l'ouverture et la mise en réseau sont sources de contraintes et de problèmes : disponibilité des équipements pour les différents utilisateurs, robustesse de l'équipement pour des utilisateurs extérieurs, mais aussi les questions de propriété intellectuelle, de confidentialité et de facturation des prestations, notamment dans le cadre d'utilisateurs privés.

Un aspect critique de l'exploitation des plateformes, soulevé par tous les intervenants, est la quotité et la pérennité des moyens humains affectés à l'équipement. Des équipements pointus, mis à la disposition d'utilisateurs variés, nécessitent absolument des personnels qualifiés, dédiés à l'exploitation de la plateforme, sa maintenance et son évolution. De plus, à côté de la gestion technique, les tâches de relation avec la communauté (formation, communication, accompagnement de projets, intégration en continu des nouveaux besoins) prennent une place croissante, décisive pour la pertinence d'une plateforme dans le temps.

Sans support ingénieur pérenne, l'obsolescence d'une plateforme est *très* rapide, la rendant caduque. Le recours à des personnels temporaires de même que la contribution de doctorants ou post-doctorants à des tâches qui ne relèvent pas de la recherche conduisent à des difficultés d'exploitation : sous-utilisation des équipements, manque de fiabilité, perte de savoir-faire, etc. En particulier, la capitalisation des connaissances est vitale pour qu'une plateforme existe dans la durée, or cette capitalisation ne peut pas se faire si elle repose seulement sur des personnels temporaires. Enfin, les plateformes en réseau posent des problèmes supplémentaires de dispersion géographique des compétences, qu'il n'est pas raisonnable de résoudre par la seule mobilité des personnels.

### **3. Propositions d'actions de l'Institut**

Il est important que l'Institut reconnaisse que les sciences de l'information ont une composante expérimentale de plus en plus grande, que l'approche expérimentale est critique dans un certain nombre de domaines si l'on veut rester à la pointe de la recherche internationale, et que des moyens substantiels doivent donc être affectés aux moyens expérimentaux. Il est également important de reconnaître que les contraintes et objectifs qui président à la construction des plateformes conduisent à des configurations extrêmement diverses, ce qui rend difficile la définition d'un modèle unique pour celles-ci. Des actions générales peuvent cependant être entreprises par l'Institut.

#### **3.1 Cartographie**

Une première étape devrait être la réalisation d'une cartographie des plateformes existantes, de leurs moyens et de leurs usages. Pour cela, une définition du périmètre considéré est nécessaire, notamment sur la taille des plateformes, ainsi qu'une description claire des buts visés par cette cartographie. L'un de ces buts devrait être de diffuser cette cartographie pour encourager des actions de mutualisation : mise en place de réseaux, partage de compétences, etc. Le but principal est l'aide au pilotage des moyens affectés aux plateformes (voir ci-dessous pour les moyens humains) et du cycle de vie des plateformes (création, utilisation selon diverses finalités, fin d'exploitation).

Une proposition de fiche est jointe à ce document pour établir cette cartographie. Elle contient d'une part des éléments factuels de description de chaque plateforme et des moyens qui lui sont affectés, d'autre part un résumé de ses objectifs scientifiques et techniques, enfin un ensemble d'indicateurs quantitatifs et qualitatifs.

### 3.2 Bonnes pratiques

Une autre action devrait viser d'une part l'aide à la valorisation des plateformes et de leurs résultats, et d'autre part l'aide à l'ouverture vers des utilisateurs extérieurs, notamment du monde industriel. Cette aide pourrait porter sur des questions de gouvernance, de management et d'évolution technique, des questions juridiques (PI, confidentialité), et des questions financières (location des équipements, accords de partenariat). Sur ce dernier point, l'évaluation, validée par l'administration, du coût complet d'une plateforme permettrait par exemple sa facturation, et la détermination d'un tarif de location.

Par ailleurs, les plateformes semblent un bon point d'entrée pour développer et mettre en œuvre les approches d'Open Science en traitant suffisamment tôt les aspects de collecte, de partage et de diffusion des données, en prenant en compte les aspects juridiques.

En pratique, il s'agirait de constituer un *guide des bonnes pratiques* couvrant notamment les aspects suivants :

- Gouvernance ;
- Moyens humains, mise en réseau des compétences ;
- Mode de fonctionnement, gestion des utilisateurs ;
- Veille technologique, évolution ;
- Evaluation du coût complet, facturation, valorisation ;
- Contribution à l'Open Science.

### 3.3 Moyens humains

L'action la plus attendue est certainement en direction des moyens humains. Il est indispensable que chaque plateforme d'envergure dispose de moyens pérennes sous forme d'au moins un ingénieur permanent affecté par l'un des organismes partenaire.

Compte tenu de l'extrême tension sur ces postes, une politique de recrutement ou d'affectation d'emplois existants devrait être affichée. La cartographie évoquée ci-dessus devrait permettre d'identifier les besoins les plus importants. L'intégration de ces personnels dans les laboratoires doit également être étudiée avec soin car les situations sont très variables. Idéalement, les personnels devraient être affectés à la plateforme et aux équipes qui les animent, pour une période donnée, et ils devraient également être rattachés à un pool technique au niveau du laboratoire ou du site afin d'assurer le bon déroulement de leur carrière.

Au-delà de cette organisation locale, l'Institut devrait encourager la constitution de réseaux de compétences permettant de mettre en relation des personnels sur l'ensemble du territoire et de mutualiser ainsi leurs connaissances et leurs expériences. Ces réseaux pourraient organiser des journées de formation et d'échanges, des visites croisées, maintenir une base documentaire et de veille technologique, etc.

Par ailleurs des actions plus ciblées, telles que des financements de CDDs, sont également bienvenues, en évitant le saupoudrage. Mais elles ne peuvent constituer la seule réponse aux enjeux des plateformes pour la recherche en sciences de l'information.

### 3.4 Chargé de mission

L'ensemble de ces actions pourrait être coordonné par un chargé de mission de l'Institut.