Des opportunités à saisir dans trois registres différents

Les explorations prospectives conduites montrent la diversité des possibles que le Big&Open Data permet d'imaginer pour le futur de la recherche. Les nombreuses ouvertures créées et l'accélération des changements qui en découlent constituent autant d'opportunités pour ceux qui sauront prendre le train en marche, et à plus forte raison pour ceux qui sauront monter dans la cabine de pilotage. Plus précisément, ces opportunités se situent dans trois registres différents, conduisant à huit défis pour la recherche publique :

- **Le développement de nouvelles capacités de recherche** grâce aux ressources numériques (données et algorithmes), qui permettent de repousser les limites de la connaissance dans de nombreux domaines et d'explorer des champs nouveaux, notamment transdisciplinaires en réponse aux questions complexes auxquelles la science doit s'attacher aujourd'hui.
- **L'émergence de nouvelles formes d'organisation** pour conduire les recherches, sortant des cadres traditionnels et associant transversalement de multiples acteurs,
- **Un renouvellement du lien avec la société**, à travers les interactions croissantes que permet le numérique entre le monde de la recherche et les citoyens, ONG et autres acteurs de la société pour la production, la qualification, la diffusion et l'utilisation de connaissances scientifiques.

Huit défis pour la recherche publique

Savoir se saisir des opportunités ouvertes par le développement du Big&Open Data est un enjeu de survie à long terme pour les institutions de recherche établies, car le train du changement ne s'arrêtera pas et le plus grand risque est bien celui de rester sur le quai. Il s'agit désormais d'innover pour faire évoluer les pratiques scientifiques elles-mêmes, en relevant un certain nombre de défis, mis en évidence par les différentes explorations prospectives :

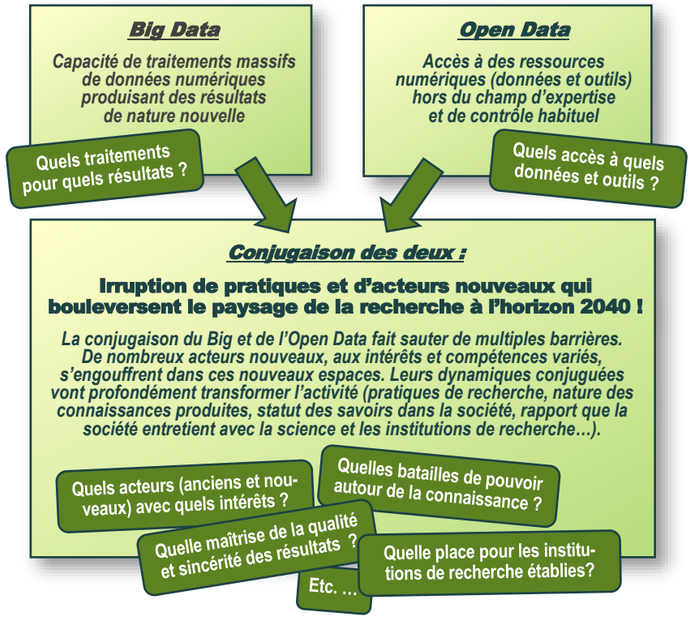
- **Repenser le positionnement des institutions publiques de recherche** dans un paysage de la recherche beaucoup plus ouvert, compétitif, rapidement évolutif, et où un nouvel équilibre est à trouver entre la production de connaissances pour le progrès des sciences et celle qui vise à satisfaire les besoins immédiats de la société.
- **Réexaminer la chaîne de valeur de la production de connaissances**, afin de mieux positionner les organismes dans un paysage concurrentiel où les interactions sont de plus en plus grandes, afin d'établir à la fois des relations saines et gagnantes avec l'ensemble des acteurs concernés et de se prémunir contre les aléas de la dépendance à des ressources numériques externes.
- **Prendre la mesure des richesses que constituent pour les organismes leurs bases de données, leurs algorithmes, leur savoir-faire**, pour les faire fructifier sans se les faire piller, et en faire un facteur d'attractivité tant par des plateformes ouvertes (mais pas forcément gratuites) du meilleur niveau que par des services innovants et expertises basées sur la recherche, associés à leur utilisation.
- **Explorer largement les alternatives aux modes traditionnels de validation et de diffusion des résultats scientifiques**, notamment en impliquant les acteurs qui donneront du crédit à ces résultats aux yeux de la société, au-delà de la qualité scientifique elle-même.
- **Réexaminer la pertinence des processus actuels d'évaluation de la recherche et des chercheurs**, dans un monde profondément transformé par le numérique, afin de promouvoir conjointement l'excellence scientifique, les approches innovantes et l'utilité sociale.
- **Savoir accompagner la diversité des modes de recherche qui se mettent spontanément en place** dans les différents écosystèmes disciplinaires, en leur offrant un cadre facilitateur et sécurisant, et en capitalisant les avantages qu'ils apportent, ceci constituant un nouveau rôle important pour les superstructures pérennes que sont les institutions de recherche.
- **Repenser l'articulation des responsabilités** entre des chercheurs de plus en plus en interaction numérique directe avec leur environnement et leurs employeurs aujourd'hui juridiquement responsables.
- **Impulser une réflexion au plan national, européen et international, sur les conditions de reconnaissance de droits d'accès et d'utilisation des données, qui seraient spécifiques à des chercheurs reconnus comme tels**, leur permettant de conduire des travaux dans des cadres déontologiques clairs et sous le contrôle d'instances appropriées, en dépassant les limitations liées aux régimes de protection des données propres à certains pays ou certains domaines.

VERSION DE TRAVAIL

Big&Open Data en recherche à l'horizon 2040

Des bouleversements majeurs, des opportunités à saisir et huit défis à relever pour la recherche publique

L'explosion des quantités de données disponibles et des moyens de les traiter d'une part (*Big Data*), l'ouverture des sources de données permettant au plus grand nombre de s'en saisir d'autre part (*Open Data*), se conjuguent pour faire des données numériques un matériau nouveau pour la recherche, accessible par chacun bien au-delà de son champ d'expérimentation et d'expertise habituel. Le monde de la recherche, dans ce qu'il est traditionnellement organisé autour de capacités d'observation et d'expérimentation par domaines disciplinaires, au sein d'institutions dédiées, se trouve interpellé et bousculé par ces changements.



Pour mieux cadrer une réflexion prospective sur un sujet aussi vaste, une réunion a été organisée par le Réseau PROSPER à la mi-juin 2017 avec les représentants d'une dizaine d'organismes de recherche publics, porteurs des principales interrogations sur l'avenir que soulevait pour leur institution la combinaison du *Big* et de l'*Open Data*.

Regroupées en sept champs de questionnement, ces interrogations ont délimité et structuré le champ d'exploration prospective, tel qu'il a été ensuite exploré par un groupe de travail constitué d'une part de porteurs de préoccupations ou d'expertises désignés par leur organisme, d'autre part de membres du Réseau PROSPER garants de la méthode. Les travaux se sont déroulés sur une année, au cours de laquelle six ateliers de travail d'une journée ont été animés par le Réseau PROSPER, qui s'est chargé de la mise en forme et de la valorisation de l'intense production de chacune de ces réunions.

Il est vite apparu que le principal accélérateur du changement était l'irruption d'acteurs nouveaux aux intérêts multiples dans le paysage de la recherche, du fait des ouvertures créées par le *Big&Open Data*. Plutôt que de s'intéresser aux détails techniques de l'évolution des pratiques de recherche liées au *Big&Open Data*, le groupe de travail a dès lors focalisé son attention sur les jeux d'acteurs et la réappropriation plus ou moins grande, plus ou moins transformante, par la société et ses différents types d'acteurs, d'une activité jusqu'alors assez protégée au sein d'un « monde de la recherche » bien organisé. Ceci explique la nature des explorations conduites (double page centrale) et des enseignements qui en ont été tirés, essentiellement tournés vers les organismes publics de recherche, à destination desquels cet exercice a été conduit (dernière page).

Ont contribué aux travaux :

***Marco BARZMAN** (INRA)**
***Philippe BIDAUD** (ONERA)*
***Arnaud BONNARD** (IFSTTAR)*
***Pascal BONNET** (CIRAD)*
***Van BUI TRAN** (IFPEN)**
***Christophe CALVIN** (CEA)*
***Patrice CARLE** (ONERA)*
***Jean-Yves CHATELIER** (INERIS)*
***David CHAVALARIAS** (CNRS)*
***Frédéric CLERC** (INRS)*
***François COSTE** (INRIA)*
***Bernard DAVID** (CEA)**
***Marie DE LATTRE-GASQUET** (CIRAD)**
***Nicolas DE MENTHIÈRE** (IRSTEA)**
***Christophe DIDIER** (INERIS)*
***Marie-Hélène EL JAMMAL** (IRSN)*
***Afonso FERREIRA** (CNRS)*
***Solenne GAUDRY** (IRSN)*
***Frédéric GOLDSCHMIDT** (IRSN)**
***Antoine GUIGON** (ONERA)**
***Pascal GUITTON** (INRIA)**
***Julien HARDELIN** (MAA)*
***Halvard HERVIEU** (MTES)**
***Emmanuelle JANNÈS-OBER** (IRSTEA)**
***Claude KIRCHNER** (INRIA)**
***Denis LACROIX** (IFREMER)**
***Philippe LEMOISSON** (CIRAD)*
***Pascale MALLARET** (IFPEN)*
***Pierre MARTIN** (CIRAD)*
***Olivier MORA** (INRA)**
***Claire NEDELLEC** (INRA)*
***Magali PENOT** (IRSN)*
***Mathieu ROCHE** (CIRAD)*
***Alain SEVE** (ONERA)*
***Maguelonne TEISSEIRE** (IRSTEA)*
***Michaël THOMAZO** (INRIA)*

** membre de l'équipe-projet PROSPER*

Réseau PROSPER

VERSION DE TRAVAIL

Sept champs d'exploration prospective sur l'impact du Big&Open Data sur la recherche d'ici 2040

Chacun des sept champs de questionnement, issus de la consultation des organismes de recherche participant à l'exercice, a fait l'objet d'une exploration prospective selon la méthode des scénarios.

Le groupe de travail a d'abord identifié les facteurs et jeux d'acteurs majeurs qui orienteront les trajectoires futures. Puis, envisageant leurs différents sens d'évolution possible, il en a déduit un éventail de scénarios, regroupés ensuite en quatre grandes tendances possibles, dont les intitulés figurent dans les bulles en périphérie de ce schéma.

Les principaux points d'attention tirés de ces explorations prospectives sont rappelés ici.

Attendus sociétaux, action publique

Le numérique, en effaçant largement les frontières et en permettant à des communautés soudées par des intérêts ou des idées communes d'agir, bouscule les notions d' « attendus sociétaux » et d' « action publique », historiquement attachées à l'Etat. Le sentiment de « faire société » devient multi-échelle, allant d'écosystèmes géographiques locaux à des réseaux globaux. Les convergences ou au contraire conflictualités mondiales, ainsi que les rapports de force ou de dialogue entre Etats, citoyens et industriels dessinent des futurs très différents pour la recherche, qui s'inscrit alors dans des cadres de compétitivité, d'éthique, d'égalité d'accès ou de finalité de la recherche très différents. Pour peser sur son avenir, le monde scientifique doit donc renforcer le dialogue tant avec les Etats qu'avec la société civile ou encore les grands acteurs du numérique.

- Cohésion et équité sociale ?
- Technoprogrès débridé ?
- Liberté socialement responsable ?
- Technoprogrès sous contrôle ?

Organisation, infrastructures, partage

Une tension constante existe à toutes les échelles (de l'Etat à chacun des acteurs concernés) entre la tentation d'autosuffisance pour sécuriser son territoire (et son activité) et le besoin d'ouverture à des ressources numériques externes. Les cadres (géo)politiques et stratégiques pèsent bien sûr dans les arbitrages au plan mondial, mais aussi des facteurs techniques comme l'interopérabilité des systèmes et langages, l'importance prise par les outils et services Big Data en ligne, ou encore le besoin de rationaliser face aux coûts croissants d'infrastructures numériques. L'échelon européen apparaît comme le bon niveau pour développer une politique coordonnée et efficace dans ce domaine, encore faudrait-il mieux articuler les stratégies des Etats et acteurs de la recherche et qu'une vraie volonté d'intégration organisationnelle de la transdisciplinarité se concrétise sur le terrain.

- Replis communautaires forcés ?
- Chacun pour soi ?
- Coopérations transnationales renforcées ?
- Nouvel ordre global ?

Accès, régulation, protection

La donnée numérique est devenue une ressource vitale pour la société. C'est donc à la fois un enjeu de pouvoir, de cohésion sociale et de progrès. Suivant que l'on est une multinationale du numérique, un Etat ou un simple citoyen, les coûts et bénéfices - ou risques et opportunités - attachés aux conditions d'accès, de régulation ou de protection des données varient profondément. Ces conditions impactent fortement les capacités productives de la recherche. Si la tendance actuelle semble plutôt à l'ouverture, la grande facilité avec laquelle les règles du jeu peuvent évoluer (pour des raisons d'ordre politique, économique ou social) conduit à s'interroger sur des dispositions spécifiques qui pourraient être négociées au niveau mondial pour permettre aux chercheurs de disposer de droits d'accès - et de devoirs associés - en toutes circonstances (adhésion à une charte d'éthique mondiale, accréditation…).

- Accès facile et usage libre ?
- Accès facile et usages contrôlés ?
- La loi des plus forts ?
- Contrôle total des usages et accès ?

Business, concurrence

Le contrôle du (ou des marchés) et de la concurrence est un enjeu qui dépasse largement les seuls aspects économiques, le Big&Open Data présentant une dimension structurante pour la société elle-même, par les valeurs qu'il contribue à construire et propager, par le langage qu'il impose et les référentiels qu'il définit. Réguler les marchés et limiter les positions dominantes est une préoccupation majeure des Etats, pendant que les consom'acteurs sont attachés à leur liberté de choix et leur besoin de confiance. Ceci permet l'émergence, à côté des marchés de masse internationaux sur lesquels les géants du numérique se livrent à une compétition féroce, de marchés plus spécifiques en réponse à des besoins précis : business de la qualité, circuits courts (pour la confiance), systèmes locaux (valeurs ajoutées spécifiques), logiciels d'exploitation libres (pour casser les mono-poles), etc.

- Les consom'acteurs ?
- Les grands prédateurs ?
- La ruche ?
- La puissance domestiquée ?

Validation/qualification, (des données, des résultats, des processus)

L'utilisation croissante de données importées à la qualité parfois difficilement maîtrisable, l'effet « boîte noire » de certains outils algorithmiques, les volumes de production qui explosent, conduisent à des trajectoires qui nécessairement renouvellent les pratiques pour assurer la qualité et la crédibilité des productions scientifiques. Dans une société construite sur l'économie de la connaissance, certains acteurs majeurs, peu intègres, pourraient confisquer à leur profit la capacité de valider ou de distribuer des données fiables. On peut aussi imaginer que la société fasse davantage confiance à des tiers pour cela, ou que des porteurs d'intérêts imposent leurs jugements sur ce que produit la science. La conception même de la validation/qualification est à repenser.

- Business avant tout ?
- A chacun sa méthode et ses finalités ?
- Maintien de méthodes éprouvées ?
- Nouvelles approches scientifiques ?

Transformation du métier

Les capacités d'investigation nouvelles offertes par le Big&Open Data relancent le besoin d'arbitrage de la société vis-à-vis de ce qu'elle attend ou demande à la recherche, entre la construction de savoirs nouveaux et la résolution de problèmes de société, entre considérations éthiques et intérêts variés. Au-delà de l'impact de nouveaux outils et ressources numériques sur les pratiques de recherche elles-mêmes (d'une simple aide à une transformation profonde), la question de la dépendance au numérique, de la maîtrise du sens de ce qui est produit ou de l'opportunité de certaines pratiques est posée. Par ailleurs, le Big&Open Data pouvant largement s'exercer hors des murs des institutions traditionnelles, la recherche devient une activité exercée par de multiples acteurs, aux compétences et aux motivations variées, plutôt qu'un métier institutionnalisé, d'où des questions nouvelles liées à la formation, la qualification et l'évaluation.

- Recherche classique ouverte ?
- Recherche opportuniste et innovante ?
- Recherche académique augmentée ?
- Recherche fondée sur le numérique ?

Connaissance produite

Qu'elles soient produites par des chercheurs ou par des machines, réservées aux élites ou diffusées dans l'ensemble de la société, les connaissances sont générées en quantités croissantes, changeant leur perception et leur usage par la société. La connaissance est de plus en plus souvent associée à l'action, produite au fil de l'eau pour répondre à des besoins, et de moins en moins d'abord considérée pour elle-même, ce qui amène à s'interroger sur la manière dont la recherche fondamentale peut être légitimée aux yeux de la société numérique du futur et soutenue par elle. Ces nouvelles finalités bousculent aussi les chaînes de valeur dans lesquelles s'inscrit l'activité des organismes de recherche. Les questions de conservation, capitalisation et réutilisation sont aussi à réexaminer, notamment au vu de la généricité des connaissances produites, c'est-à-dire leur aptitude à être utiles dans de larges champs d'application.

- La mitraille algorithmique ?
- L'intelligence des machines ?
- Les savoirs étendus des experts ?
- Les nouvelles connaissances de la société ?