

Big Data Feuille de route

La feuille de route a été présentée le 2 juillet 2014 par MM. Paul HERMELIN (PDG de CapGemini) et François BOURDONCLE (Président de FB&Cie), chefs de projet, au comité de pilotage des plans de la Nouvelle France Industrielle.

L'équipe projet

Présidents : François Bourdoncle (Président de FB&Cie), Paul Hermelin (Président Directeur Général de Capgemini),

Membres industriels : Vivek Badrinath (DGA Accor), Marc Chemin (Capemini), Jérôme Gueydan (Orange), Jamal Labed (AFDEL), Romain Niccoli (Criteo)

Correspondants cabinet MEIN : Bertrand Pailhès, Bruno Sportisse

Correspondants Direction Générale des Entreprises (DGE) : Fabien Terrailot, Cédric Mora

Correspondants des autres ministères : Laurent Rojey (CGI – Commissariat Général à l'Investissement)

Les objectifs du plan

Le Big Data est souvent décrit comme l'ensemble des technologies permettant la valorisation des données massives, mais c'est aussi, et surtout, une « arme de guerre » des grandes sociétés du numérique pour s'attaquer méthodiquement à des pans entiers de l'économie (industrie et services) qui n'ont pas encore été impactés par la révolution numérique. A l'instar de ce qui s'est passé pour la presse, le commerce de proximité, le tourisme ou encore l'industrie musicale, ces sociétés tentent actuellement d'utiliser :

- les centaines de millions de clients avec lesquels elles ont une relation de proximité et de confiance,
- les nombreuses données qu'elles ont sur ces clients,
- et leur capacité de traitement et de croisement de ces données pour modifier profondément le paysage concurrentiel, et en particulier la relation client et la chaîne de valeur.

C'est ainsi que des secteurs comme l'assurance, la banque, la grande distribution, le crédit à la consommation, l'industrie automobile, les « utilities », l'énergie, et bien d'autres encore risquent de voir leur positionnement dans la chaîne de valeur réduite progressivement au rôle de sous-traitant « technique », avec une très forte pression sur les marges et un risque potentiellement létal pour les plus faibles.

Les activités liées au Big Data ont représenté en France environ 1,5 milliards d'euros cette année, et atteindront près de 9 milliards d'ici 2020 (recouvrant les dépenses en logiciels, services, et dépenses internes des entreprises). En termes d'emplois, l'enjeu du Big Data est de créer ou consolider de l'ordre de 137 000 emplois, soit directement dans l'industrie informatique, soit dans les fonctions technologiques au sein des entreprises, soit dans les fonctions métiers (par exemple les fonctions de vente) des entreprises.

L'enjeu disruptif potentiel sur l'économie française et un grand nombre de ses industries et services est tout à fait considérable. Et le Big Data reposant sur la *réutilisation à des fins différentes* et le *croisement* de données existantes ou nouvelles, il résulte que le développement du Big Data va concerner un nombre croissant de sociétés et que ce développement va se heurter très directement au principe de « respect de la finalité initiale de la collecte des données » qui sous-tend la loi « Informatique et Libertés ».

Sur la base de ce constat, la présente feuille de route propose que le Plan Industriel Big Data s'articule autour d'actions spécifiques visant à lever des verrous ou à développer le marché du Big Data par une politique volontariste de l'offre.

➤ **Mesures concernant l'écosystème**

L'ensemble des mesures proposées vise à renforcer l'offre de technologies et/ou services Big Data.

1. **Formation**. S'assurer que notre système éducatif forme suffisamment de « data scientists », profil qui va être au cœur de cette nouvelle révolution industrielle. Les formations devront être à la fois des formations courtes, de type formation continue ou professionnalisante, des mastères spécialisés, et des formations d'ingénieur de niveau bac +5 plus classiques.
2. **Accès des start-ups aux données et aux infrastructures**. Favoriser l'innovation ouverte, élément déterminant dans une stratégie globale de réponse de nos grands groupes, en promouvant le développement des startups dans le domaine du Big Data via la mise en place d'un « Centre de Ressources Technologiques » fournissant :
 - Recensement des données sur lesquelles les grands donneurs d'ordre sont prêts à travailler dans une logique d'innovation ouverte (grands groupe, services de l'Etat, startups), et garantie de la neutralité dans l'accès à ces données (sans que ce ne soit nécessairement des données « ouvertes »)
 - Consolidation des initiatives privées de plateforme de collecte de données
 - Accès à des compétences et des logiciels spécialisés dans le domaine du Big Data
 - Accès à des ressources de calcul à grande échelle pour mener des expérimentations réalistes et construire rapidement des prototypes opérationnels
3. **Soutien à l'accélération des start-ups** Soutien à la création d'accélérateurs et de fonds d'amorçage spécialisés dans le domaine du Big Data (accompagnement par des entrepreneurs de la croissance de start-up)
4. **Observatoire des usages Big Data**. Observatoires des usages, identification et soutien des nouveaux projets qui émergeront dans la sphère Big Data.

➤ **Mesures sectorielles :**

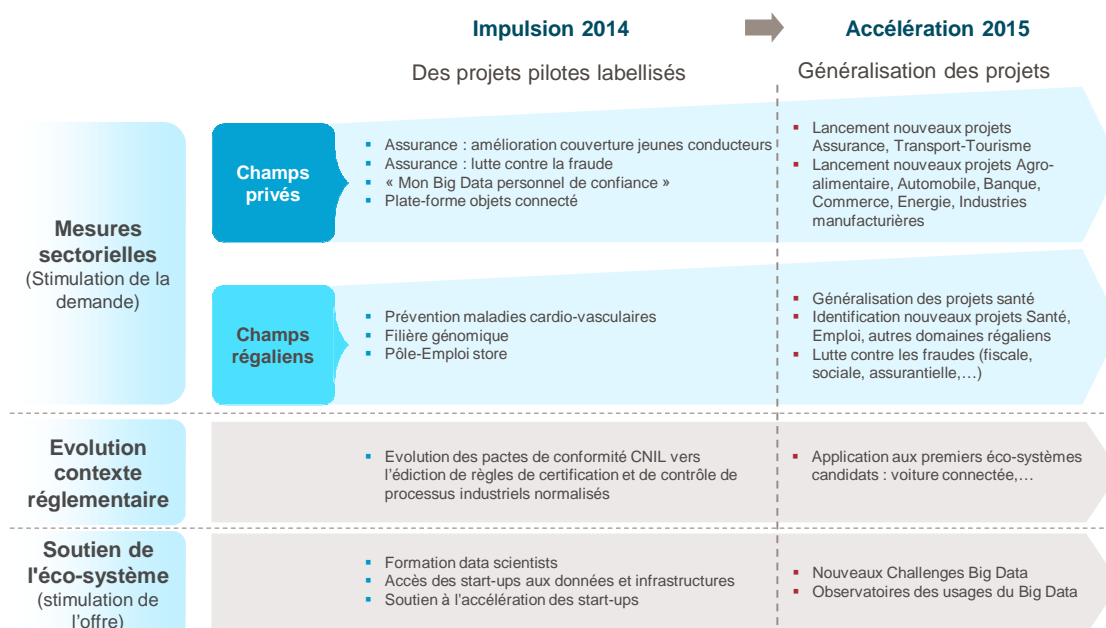
Les actions proposées dans cette section sont des projets concrets, portés par des acteurs industriels (ou des opérateurs dans le cas du secteur public), qui visent à accélérer le développement de la demande, i.e. des usages du Big Data.

5. **Projets sectoriels dans le secteur privé :** Identifier les secteurs industriels les plus à risque dans le domaine concurrentiel (Assurance, Energie, Transport) et permettre à ces acteurs privés de mener à bien des expérimentations sur de nouveaux usages de la donnée leur permettant de mener une politique industrielle offensive sur le plan international dans le domaine.
6. **Projets sectoriels dans le secteur public :** Faire de l'Etat un exemple dans le domaine du Big Data, en lançant de grands projets dans les domaines de 1 / la Santé, 2/ l'Emploi. Ces grands projets doivent être l'occasion de faire rentrer dans une logique d'innovation ouverte les services de l'Etat concernés.

➤ **Mesures concernant le contexte réglementaire et législatif**

7. **Réglementation et législation.** Faire évoluer la mise en œuvre de la Loi Informatique et Libertés et la CNIL vers une logique de « certification de processus » et de « label éthique » permettant aux industriels qui y auront recours de promouvoir un approche éthique de l'utilisation des données personnelles et d'en faire un atout concurrentiel, notamment à l'exportation. Le plan s'est attaché à définir des actions concrètes avec la CNIL à cadre législatif constant, partant notamment de la démarche de « Pacte de Conformité » déjà initiée par la CNIL.

Synoptique de la feuille de route



1

Liste des actions

1. Actions de nature à développer l'écosystème Big Data en France	
<i>Action 1</i>	Formation de « data scientists »
<i>Action 2</i>	Accès des start-up aux données et aux infrastructures
<i>Action 3</i>	Soutien à l'accélération et au financement des start-up Big Data
<i>Action 4</i>	Observatoire des usages du Big Data
Actions de nature à mettre en avant des initiatives sectorielles Big Data en France	
<i>Actions 5</i>	Projets sectoriels dans le secteur privé
<i>5.1</i>	Assurance
<i>5.2</i>	Création d'un espace de confiance personnel – « Personal Big Data »
<i>5.3</i>	Création d'une plate-forme ouverte de données des objets connectés
<i>5.4</i>	Smart City energy analytics
<i>Actions 6</i>	Projets sectoriels dans le secteur public
<i>6.1</i>	Secteur de l'Emploi
<i>6.2</i>	Secteur de la Santé - Programme de sensibilisation, de repérage et d'accompagnement des assurés MGEN portant sur la sphère cardio-vasculaire
<i>6.3</i>	Filière génomique : séquençage du génome et oncologie
Actions de nature à modifier la réglementation autour du Big Data en France	
<i>Action 7</i>	Evolution du contexte réglementaire et législatif

Actions de nature à développer l'écosystème Big Data en France

Le Big Data est un domaine qui va susciter une demande rapidement importante de profils connus sous le nom de « data scientist », à même d'appréhender à la fois les méthodes statistiques, mathématiques et informatiques dans le cadre d'un contexte métier particulier. Le système éducatif français doit donc s'adapter en conséquence.

D'autre part, le fort impact de la révolution du Big Data pour toutes les industries et les services traditionnels va nécessiter une mobilisation de toutes les énergies et imposer plus que jamais aux grandes sociétés de travailler dans une logique d'innovation ouverte pour s'adjoindre des métiers nouveaux leur permettant de faire évoluer leur relation client dans la bonne direction. C'est la raison pour laquelle le soutien à l'écosystème innovant via des fonds d'amorçage dédiés est très important, ainsi que la création de centres permettant aux startups d'accéder facilement à des jeux de données, des compétences et des moyens de calcul leur permettant de développer très rapidement des innovations de rupture, condition indispensable pour espérer devenir à brève échéance des leaders mondiaux.

Action 1 : Formation de « data scientists »

Objectifs et motivations

La « science des données » (data science) est la science qui sous-tend l'ensemble des technologies du « Big Data ». La France disposant d'excellents enseignements en mathématiques, statistiques et informatique, elle a le potentiel de former des data scientists au meilleur niveau international.

Que faut-il enseigner ?

Le consensus aujourd'hui est de définir le data scientist à l'intersection de trois domaines d'expertise : (i) l'informatique, (ii) les statistiques et les mathématiques, et (iii) les connaissances métier. En formation initiale, il sera pertinent d'intégrer des séjours longs en entreprise pour que les futurs diplômés maîtrisent le contenu sectoriel des métiers. En formation continue, l'accent portera avant tout sur une meilleure maîtrise de l'informatique et des méthodes de traitement des données.

Si pour une partie des futurs professionnels de ce secteur, il est nécessaire de concevoir des formations complètes qui permettent d'identifier dès la formation initiale des compétences indispensables aux professionnels de ce secteur, le profil de ces nouveaux professionnels doit être déterminé dans toute sa variété et avec l'ensemble des niveaux de formation à identifier.

Les métiers périphériques et partenaires des « data scientist » constituent aussi une cible de formation, pour lesquels une part de différenciation doit garantir une interaction optimale avec les spécialistes du secteur.

La question de la transformation et de l'évolution des compétences de professionnels actuellement à la périphérie de ce secteur est aussi un axe qui doit permettre de façon plus rapide d'avoir des professionnels « data scientists » qui auront la capacité d'amener avec leur expérience et expertise précédente des modalités professionnelles variées (statistiques, informatique, SI, ...)

Des projets « les yeux dans les yeux de données en vraie grandeur » complèteront ces cours en faisant intervenir tous les aspects de la science des données, depuis l'acquisition jusqu'au plan d'exploitation. Ces projets doivent encourager **la créativité** des élèves et s'appuyer, en particulier dans le cadre de la formation continue, sur leurs compétences existantes, leurs intérêts, leur domaine d'expertise. Ces formations s'appuieront sur des outils logiciels qui pourront être présentés par des industriels du domaine et utilisés pour les expérimentations. Une telle formation est par nature pluridisciplinaire et il est indispensable de ne pas occulter une de ses dimensions techniques ou sa dimension métier :

- **Complétude.** Nous parlerons ici de formation « complète » pour les formations tenant compte de ces aspects dans leur totalité. Si nous insistons sur ce point, c'est qu'aujourd'hui, par effet de mode, on fait parfois le marketing de formations informatiques ou de formations de statistiques comme formation de data scientists. Une telle approche est positive en ce qu'elle permet de faire bouger les lignes, il ne faut pas que cela perdure. La formation doit être à terme pluridisciplinaire. Selon les organismes formateurs, on pourra avoir cependant sans doute, à terme, des formations « à majeure » informatique, statistique ou métier.
- **Différentiation tardive.** La formation doit offrir un tronc commun le plus généraliste possible. La raison est simple : il est simplement impossible d'offrir des formations de data scientists spécialisées à chacun des métiers concernés. La formation de data scientists devra par contre offrir des spécialisations métiers en fin de cursus peut-être par le biais d'options, certainement à l'occasion de projets avec des interactions fortes avec les entreprises.
- **Illustration.** On a de tout temps réalisé la difficulté de faire comprendre les statistiques aux décideurs et au grand public. C'est bien là tout le spectre d'un data scientist.

Actions proposées

Actions à court terme :

1. **Formations en ligne.** Il faut développer, dans ce domaine, les Moocs (Massive on-line open courses) ou autres formations à distance en insistant sur la qualité et la complétude. De tels cours peuvent être déployés très vite. Un but serait notamment de rendre disponible assez vite des cours qui servent de référence dans le monde francophone. Pour y arriver, on pourra s'appuyer sur la plateforme Fun du MENESR ; il faudra affecter de vraies ressources : temps de préparation du cours, TP, plateforme technique (matériel, logiciel et jeux de données), support technique. Par exemple, l'ordre de grandeur pour produire un vrai Mooc est au minimum l'équivalent d'un professeur temps plein sur un an. On s'appuiera sur des cours existants de qualité pour développer une offre de cours dématérialisés. Ce type de formation pourra servir à la formation continue dans les entreprises et à la formation des enseignants eux-mêmes.
2. **Curriculum et label « data science ».** On définira précisément un curriculum général de formation de data scientists ainsi que de formations spécialisées dans des domaines comme le marketing, les sciences, le journalisme... Les formations continues pourront bénéficier d'un label qualité, basé sur un référentiel de compétences à développer.
3. **Formations professionnalisantes.** On soutiendra leur développement en insistant sur la qualité et la complétude :

- a. Stages courts de quelques jours (établissements d'enseignement supérieur, entreprises de formation)
 - b. Formations diplômantes (établissements d'enseignement supérieur)
- 4. Faire évoluer les formations actuelles**
- a. Encourager les formations statistiques à inclure plus d'informatique et les formations informatiques à inclure plus de statistique, dans une optique d'acquisition de compétences mixtes. .
 - b. Encourager à introduire une vraie exposition aux aspects business
 - c. Identifier des intervenants universitaires ou industriels pour combler les manques.
5. **Réseau.** On associera les entreprises à la formation de data scientists par le biais d'un réseau avec une communauté de partage de ressources (outils, données, formation) et d'expériences. Ce réseau pourra en particulier proposer des projets d'expérimentation aux apprenants. Il pourra aller jusqu'à la mise en place de collaborations. Ce réseau pourrait se faire dans le cadre du Centre de ressource technologique proposé dans le cadre de la Filière Big data et en étroite interaction avec les conseils de perfectionnement des formations de l'enseignement supérieur..
6. **Nouveaux centres d'éducation.** Envisager des coopérations avec des écoles comme Simplon.co ou 42 et des expériences nouvelles d'éducation des data scientists.

Pour structurer des formations à plus long terme, les actions sont les suivantes :

1. **Evaluation** : Réaliser de manière prioritaire et urgente une évaluation des besoins tant qualitatifs que quantitatifs en lien notamment avec les branches professionnelles.
2. **Les formations métiers.** Développer la formation de data scientists dans les formations métier : universités, écoles d'ingénieurs, école de management et de commerce, école de journalisme, école d'administration.
3. **Le droit des données.** Développer dans les études juridiques la place des droits du numérique, et en particulier des droits des données.
4. **Matériel éducatif** : développer une gamme de matériels éducatifs ouverts et de qualité, appuyés sur le référentiel de compétences, comme des livres, des collections de données tests, des Moocs.
5. **Assouplissement des structures éducatives publiques.** Les formations dispensées par des établissements d'enseignement supérieur habilités par l'état sont évaluées très régulièrement. Ainsi, leurs évolutions sont soumises à des conseils de perfectionnement qui impliquent très généralement des entreprises. En outre, la mise en place de l'accréditation permet plus rapidement au sein d'une mention existante de transformer un parcours ou d'en créer un à la seule autonomie des établissements. La souplesse de ces procédures pourra être ajustée afin de répondre aux besoins du domaine des Big Data qui évolue très vite. Les conventions de stages et les facilités de mobilité entre entreprises et établissements d'enseignement supérieur pourraient être aménagées de manière à être plus facile dans leur mise en œuvre. La formation continue publique fonctionne plutôt bien en ce qui concerne la formation diplômante. La culture de la formation à la demande (stages intensifs) doit être généralisée.
6. **Bac+3.** Les textes d'applications de la loi sur le cadre national de formation permettent de mettre en place une spécialisation progressive, qui permet d'envisager un parcours de licence bi-disciplinaire Mathématiques-Informatique orienté autour des données et permettant une formation équilibrée qui pourrait d'une part satisfaire les besoins des entreprises à ce niveau-là et servir de socle à des Masters en Sciences de Données.

Action	Formation de « data scientists »
Pilote	MENESR
Acteurs concernés privés	Branches professionnelles
Levier(s) public(s) à actionner	Besoins de soutiens spécifiques ?
Acteurs concernés publics	Universités et Grandes Ecoles, en allant du bac +2 (BTS, DUT...) au doctorat (par exemple : Institut Mines Telecom, ENSAE, ENSAI, Ecole Polytechnique, Grenoble INP et Ecole de Management, ESSEC, HEC...) Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche
Calendrier de mise en œuvre	Septembre 2014 : mise en place de MOOC Novembre 2014 : recensement des formations existantes

Action 2 : Accès des start-up aux données et aux infrastructures

Objectifs et motivations

Contrairement à la précédente phase d'innovation IT, le Web 2.0, où il était possible de créer un nouveau service Internet (réseau social par exemple) en isolation dans un garage avec son PC, le Big Data présente à l'inverse trois verrous simultanés :

1. l'accès à de grands volumes de données à analyser pour fournir le service, dont l'innovateur ne dispose jamais au début ;
2. la maîtrise de technologies d'analyse de données et de « machine learning » très complexes et en constante évolution ;
3. l'accès à des ressources de calcul très puissantes pour exécuter les algorithmes d'analyse de données à très grande échelle, ce qui est difficile pour une startup même si les systèmes « à la demande » disponibles dans le Cloud commencent à simplifier ce problème.

Le premier point est particulièrement préoccupant, car l'écosystème de l'innovation actuel, copiant en cela ce qui se fait aux Etats-Unis, est en train de développer un modèle d'innovation intégré « fonds d'amorçage + accélérateur ». Autant ce modèle fait du sens pour les méga-fonds de capital-risque américains qui ont la masse critique, le risque pour l'écosystème français est d'avoir toute une série d'accélérateurs sous-critiques qui, pour se différencier sur le marché, vont être tentés (et ils le font déjà) de faire des partenariats privilégiés avec des grands groupes pour fournir aux start-up de leur programme les données que ces sociétés souhaitent soumettre à l'innovation ouverte. Un entrepreneur souhaitant par exemple travailler dans le domaine de la relation client des opérateurs télécom va donc devoir faire très tôt un choix cornélien entre plusieurs accélérateurs travaillant de manière exclusive avec des acteurs établis ; et choisir au passage le fonds d'amorçage qui lui est associé. Cela nous semble très dangereux, car une fois ce choix fait, il sera très difficile pour cette start-up de travailler avec les autres opérateurs, et cela compromettra gravement ses chances de succès. A l'inverse, le modèle du « grand accélérateur avec grand fond d'amorçage intégré » n'a aucun sens, car l'écosystème dispose déjà de très nombreux accélérateurs et il est absurde de vouloir aller contre cette tendance de fond.

Si notre pays souhaite faire émerger de nouveaux leaders mondiaux dans le domaine du Big Data, il est critique de se concentrer sur ces trois points de manière à pouvoir réduire au maximum le « time to market » de nos start-up et maximiser leurs chances d'être premier entrant sur leur marché et de conquérir le marché mondial avant d'être copiées ou rachetées par des entreprises américaines notamment.

Action proposée :

- Création des conditions d'un ou plusieurs centres, qui seront un lieu privilégié pour les acteurs privés ou publics, détenteurs de jeux de données, souhaitant travailler avec des start-up dans un modèle « d'innovation ouverte ». Ces centres permettront de créer des partenariats dans la durée entre des entreprises possédant des données massives, et des start-up ayant de fortes capacités d'innovation sur ces sujets, et ce dans environnement

adapté (échanges privilégiés, mise à disposition de jeux données, accords de confidentialité, etc.). De plus, des moyens humains doivent être mis à disposition de ces start-up pour leur permettre de développer rapidement les compétences nécessaires à la maîtrise de l’algorithmique spécifique du domaine. Enfin, des infrastructures de calcul devront être proposées dans ce cadre pour laisser les start-up effectuer très rapidement des travaux à échelle sur leurs données, au moins d’un point de vue expérimental, de manière à pouvoir tester leur innovation et lever de l’argent rapidement dans de bonnes conditions. Ces infrastructures pourront s’appuyer sur des solutions proposées dans le monde de la recherche ou via des partenaires commerciaux proposant ce type de plateforme.

Action	Accès des start-up aux données et aux infrastructures
Pilote	CGI / DGE au titre de l’Appel à Manifestation d’Intérêt « Challenges Big Data » du programme des Investissements d’Avenir
Acteurs concernés privés	Acteurs susceptibles de répondre à l’Appel à Manifestation d’Intérêt
Levier(s) public(s) à actionner	Appel à Manifestation d’Intérêt « Challenges Big Data »
Acteurs concernés publics	BPI

Action 3 : Soutien à l'accélération des start-up

Le caractère disruptif de l'innovation introduite par le Big Data dans un champ toujours plus large de domaines industriels ouvre d'importantes opportunités de création de valeur et de différenciation, notamment pour des startups et nouveaux entrants fournisseurs de technologies.

La capacité pour ces entreprises à transformer de façon pertinente d'énormes quantités d'informations en solutions innovantes et nouveaux modèles économiques, constitue un enjeu très fort. Les startups ont la capacité d'innover rapidement pour tester et proposer de nouveaux services autour du Big Data et ont donc un avantage compétitif sur le marché – d'autant plus si elles créent et accumulent des données en propres – et peuvent capter une part essentielle de la valeur.

La création ou le renforcement d'entreprises fournisseurs de technologies ou innovant dans les usages du Big Data et capables de saisir les opportunités économiques du Big Data, nécessite un accès à des financements adaptés. Une intervention de l'État pour soutenir le développement des start-up du domaine, souvent très consommateur de capital dans les premières phases, permettrait d'accélérer la création de champions technologiques au niveau français et international.

Les accélérateurs d'entreprises ont aussi un rôle majeur dans le soutien au développement des startups du secteur. L'Etat regardera très attentivement tous les projets de création d'accélérateurs proposant une aide aux entreprises exploitant ces données massives et souhaitant s'intégrer dans l'initiative French Tech.

Action proposée :

- Lancement de défis de valorisation des données massives dans le cadre du Concours Mondial d'Innovation suite au rapport de la commission "Innovation 2030". Ce concours permet de sélectionner des projets au stade amont de leur développement (amorçage) ou pour des travaux de développement de plus grande ampleur (accompagnement), et de leur apporter une aide financière.
- Abondement de fonds d'amorçage spécialisés dans les technologies du Big Data ou de la transformation numérique des entreprises liée aux nouveaux usages du Big Data
- Soutien aux éventuels projets d'accélérateurs Big Data qui pourraient émerger dans le cadre de l'initiative French Tech.

Action	Soutien au financement et à l'accélération des start-up
Pilote	BPI
Acteurs concernés privés	Accélérateurs French Tech et fonds d'amorçage
Levier(s) public(s) à actionner	Concours Mondial d'Innovation – Thématique Big Data
Acteurs concernés publics	MEIN, DGRI, CGI, DGE
Calendrier de mise en œuvre	2014 : Financement pour les startups et PME dans le cadre du Concours Mondial d'Innovation (phase d'amorçage et d'accompagnement)

Action 4 : Observatoire des usages Big Data

L'observatoire Big Data permettra de continuer, dans la durée, l'identification et le suivi des nouvelles initiatives en lien avec le Big Data.

Le groupe de travail recommande de maintenir un dialogue permanent avec l'ensemble de l'écosystème public et privé, pour :

- Continuer « d'évangéliser » les acteurs économiques sur la nécessité d'adopter des solutions Big Data pour adapter leurs produits et services
- Identifier (et soutenir le cas échéant) des nouveaux projets Big Data arrivant à maturation
- Suivre et quantifier les enjeux liés au Big Data, et notamment les emplois et compétences

Des actions pourront être menées à différents niveaux, comme (liste non exhaustive) :

- L'organisation de rencontres semestrielles entre les ministres et des PDG des grandes entreprises des secteurs concernés par l'irruption de nouveaux modèles permis par le Big Data (par exemple : assureurs, ...)
- L'organisation avec la BPI de rencontres entre grandes entreprises et start-up
- La quantification des enjeux économiques et des emplois induits, (y compris en intégrant les conclusions du rapport de la mission Lemoine lorsqu'elles auront été remises)

Pour ce faire, une action de communication et d'interaction régulière entre le Ministère de l'Economie, de l'Industrie et du Numérique, et les acteurs économiques de l'ensemble des secteurs d'activité sera animée pour maintenir la dynamique du plan.

Action	Identification et suivi de nouvelles initiatives
Pilote	DGE
Acteurs concernés privés	-
Levier(s) public(s) à actionner	Aucun
Acteurs concernés publics	DGE, BPI, associations professionnelles (AFDEL, ...)

Actions de nature à mettre en avant des initiatives sectorielles Big Data en France

Action 5.1 : Assurance

Objectifs et motivations

La quantité de données a toujours été l'élément majeur de la fiabilité des statistiques, fondement de l'assurance. Quantité, mais aussi qualité des données, sont par conséquent des facteurs essentiels de différenciation compétitive pour les assureurs.

Les compagnies d'assurance, tout comme les banques ou les opérateurs télécoms, disposent d'un grand nombre d'informations sur leurs clients, données relativement peu exploitées aujourd'hui.

De plus, ces données internes peuvent encore être enrichies par des données externes :

- les données open data. Exemple : l'INSEE met à disposition des données socio-démo consolidées au niveau du quartier, permettant de probabiliser des profils de clients en fonction de leurs adresses
- les données de comportement sur les réseaux sociaux et sur le web. Exemple : de plus en plus d'acteurs utilisent des données du profil Facebook (sur base d'opt-in) pour affiner le profil de leurs clients
- les données issues des objets connectés, qui vont se développer. Exemple : le développement des voitures connectées permettra de collecter des informations sur le comportement des automobilistes

Le développement de nouvelles technologies dans la sphère du big data et des analytics permet désormais de mieux stocker et analyser ces données, avec de nouvelles opportunités pour le secteur de l'assurance à la clé.

L'exploitation du big data par l'assureur recherchera pour l'essentiel, l'amélioration de l'efficacité commerciale. Pour autant, chaque étape du cycle de vie de gestion du client assurantiel sera impactée. Il est à noter que ces bénéfices du big data pourront également profiter à l'assuré comme à la société au global.

Exemples d'applications possibles du big data dans le domaine de l'assurance :

1. **Acquisition de client** : meilleure gestion des campagnes publicitaires et commerciales et connaissance de la sensibilité des clients à ces actions, analyse des canaux de distribution à des fins de redéploiement ou de segmentation, optimisation des coûts d'acquisition des nouveaux clients.
2. **Souscription** : sans doute le maillon de la chaîne de valeur qui sera le plus impacté avec la mise en place d'une « nouvelle science du risque », où la qualité de la prédiction s'appuiera sur une quantité d'informations disponibles considérable et permettra une sélection et une tarification véritablement personnalisée.
3. **La gestion client** : connaissance client enrichie, ce qui va créer des possibilités considérables en matière de segmentation, développement de nouveaux produits, génération de cross et up selling aux moments clés de la vie, prévention des résiliations et

simulation des ressorts de la fidélité client, etc. Le lien client va s'en trouver renforcé, car les assureurs pourront améliorer la gestion individuelle via une vue temps réel, omni canal des clients, mieux communiquer avec eux et véritablement proposer la « next best action », à savoir le contact, l'information, le produit ou le service personnalisé pertinent à l'instant t.

4. **La gestion de sinistres** : possibilité de pousser des messages de prévention « ciblés » et à l'efficacité renforcée, mise en place de modèles de prévisions de sinistres dommages mais aussi de façon plus large des crises sanitaires, catastrophes naturelles, risques environnementaux, identification et limitation des fraudes à l'assurance, amélioration des systèmes de transport et de santé, etc.
5. **Nouveaux usages connectés** : de la capacité qu'auront les assureurs de développer de nouveaux usages connectés (Quantified Self, IoT, services à la personne, surveillance médicale à distance, etc.) dépendra leur capacité à analyser finement les usages, détecter plus finement les bons ou les mauvais clients, faire des offres de produits innovants et au juste prix, et à terme, maintenir leurs marges face aux nouveaux produits d'assurance qui pourraient être proposés par les géants du numérique.

Si le big data est un formidable outil de connaissance du client, la mise en place d'analyses prédictives qu'il autorise, peut poser problème.

En effet, la richesse de l'information pourra conduire l'assureur à développer une sélection forte et détruire le principe même de mutualisation du risque, socle de l'assurance depuis les origines.

De la même manière, les assurés pourront également mieux choisir leurs contrats, en fonction de besoins mieux connus de part et d'autre. Les assureurs risquent ainsi de voir les « bons risques » leur échapper ou bien de devoir les assurer à très bas prix.

Il appartiendra donc aux compagnies d'assurance de redéfinir les règles de mutualisation pour gérer l'équilibre des risques tout en tirant partie des possibilités de segmentation fine du big data.

Pour exploiter au mieux ce nouveau potentiel du big data, on voit que les compagnies d'assurance vont devoir faire évoluer leur façon de travailler. De façon à initier et accélérer cette transformation, nous recommandons de commencer par des PoC (Proof of Concept) sur de premiers périmètres de données, qui permettront d'une part de prouver les bénéfices apportés par l'utilisation des analytics, d'autre part de mettre en évidence les moyens à mettre en place.

Différentes initiatives seront soutenues dans le secteur de l'assurance.

Actions proposées :

Ces expérimentations (PoC) pourront donner lieu à des projets sur le court terme comme la définition de polices d'assurance plus adaptées aux profils des assurés porté ici par Covéa.

On peut aussi citer les projets suivants :

Donner la possibilité aux assurances de valoriser et monétiser le patrimoine de données qu'elles ont accumulées depuis des années.

Square Predict est un projet open source qui a pour objectif de bâtir une plateforme SaaS permettant aux assureurs de réaliser des prédictions à partir des données de leurs assurés croisées avec celles disponibles sur Internet (réseaux sociaux, web sémantique, open data...). Les partenaires sont : un géant de l'assurance, trois laboratoires de recherche (LIPN, LIPADE et LARIS), deux éditeurs et un expert en protection de la vie privée,

L'enjeu pour les producteurs d'assurance est d'arriver à exploiter cette avalanche de données pour ne pas se voir concurrencés dans la distribution de leurs produits par des géants américains qui ont développé une réelle expertise dans la fouille de données. En effet, la capacité à proposer des produits ciblés, en traitant les immenses volumes de données en leur possession donne un avantage concurrentiel sans égal à ces grands acteurs de l'Internet. Square Predict est donc un projet d'analyse décisionnelle visant à valoriser et monétiser l'énorme quantité de données des assureurs.

Square Predict pourrait ainsi résoudre les problématiques suivantes :

- Estimation en temps réel de l'impact, y compris financier, des dégâts causés par un événement climatique de grande ampleur ;
- Prédiction de la future sinistralité incendie/dégâts des eaux en fonction des données publiques accessibles (quartier par quartier, voire rue par rue, voire adresse par adresse) et de l'historique des assurés ;
- Détection de l'apparition d'événements globaux (comme la grippe) par l'analyse des signaux faibles.

Square Predict est un projet.

Challenges Big Data dans le domaine de l'assurance

Cap Digital, en tant que Pôle de Compétitivité, est un des animateurs sélectionné pour organiser des challenges Big Data (voir Action 2).

La thématique de l'Assurance, via les sponsors identifiés, représente un des premiers sujets qui sera proposé dans ces challenges Big Data.

Action	Projets dans le domaine de l'assurance
Pilote	Covéa, Arrow Group, Cap Digital
Acteurs privés concernés	Covéa, Arrow Group, Isthma (groupe Viseo), Digital & Ethics
Levier(s) public(s) à actionner	Appels à projets Big Data, Appel à Manifestation d'Intérêt « Cœur de filière numérique – challenges Big Data ».
Acteurs publics concernés	DGE, BPI, CNIL
Calendrier de mise en œuvre	Lancement en 2015

Action 5.2 : Création d'un espace de confiance personnel : projet Mes Infos

L'immense majorité des projets "Big Data" en explore le potentiel *au service des organisations*. Ce projet se propose d'explorer dans quelle mesure le Big Data pourrait apporter *aux individus* plus de pouvoir sur leurs propres données, plus de connaissances et de maîtrise de leur vie.

Plusieurs raisons pressantes invitent à explorer cette voie nouvelle :

- Les individus ont désormais accès à des capacités informatiques considérables sur leurs smartphones, leurs ordinateurs et le *cloud*.
- Les deux décennies passées montrent que l'innovation naît au moins aussi souvent sur les marchés grand public que sur les marchés "B2B".
- Alors que la société se montre de plus en plus sensible à la protection des données personnelles, les bénéfices du Big Data ne peuvent pas être entièrement accaparés par les organisations, au risque de produire une grave crise de confiance. En l'état, les "3V" du Big Data (volume, variété, vélocité) se heurtent en effet aux grands principes des législations européennes en matière de données personnelles (finalité, proportionnalité, durée limitée de conservation, transparence...)
- A l'inverse, des études ou des services innovants exigeant le croisement de nombreuses données hétérogènes ne seront réalisables qu'à l'initiative et sous le contrôle des individus, qui doivent par conséquent en tirer bénéfice.
- Enfin, une approche décentralisée, qui part des individus, constitue la meilleure chance d'éviter la domination massive du marché des Big Data par quelques plates-formes géantes (les "GAFA"), principalement américaines.

Le projet MesInfos : mettre le pouvoir des données à portée des individus

Depuis 2 ans, sous l'impulsion de la Fing (Fédération Internet Nouvelle Génération), une communauté de grandes entreprises¹, chercheurs, entrepreneurs et designers, avec le concours de la plate-forme de "cloud personnel" Cozy Cloud, explore ce qu'il se passerait si les individus disposaient des moyens de recevoir et produire, puis d'exploiter leurs données personnelles, sous leur contrôle et à leurs propres fins. Il est en effet possible d'imaginer de nombreux usages du "Big Data" au service des individus :

- Mieux se connaître soi-même, revenir sur ses expériences, se comparer (domaines d'usage : finances ; forme et santé ; développement professionnel ; énergie...)
- Prendre de meilleurs décisions, comparer des offres complexes, exprimer ses intentions sur des places de marché numériques (usages : "*vendor relationship management*" ou VRM ; gestion d'événements de vie ; mobilité ; choix financiers...)
- Contribuer volontairement à des analyses massives à l'aide de données anonymisées (projets urbains ; études fondées sur des données sensibles : études cliniques...)

Cette piste n'est à l'heure actuelle explorée que dans deux autres pays : les Etats-Unis (*Green Button*, *Blue Button*, communauté VRM) et le Royaume-Uni (Midata). La France fait ainsi partie

¹ AXA, La Banque Postale, Crédit Coopératif, Société Générale, GS1, Intermarché/les Mousquetaires, Orange, Solocal Group. Tous ces partenaires ont confirmé leur engagement dans cette nouvelle étape du projet.

des pionniers, d'autant que MesInfos vient de conclure une phase d'expérimentation de terrain qui n'a pas d'équivalent ailleurs.

Vers le "Personal Big Data"

En 2014-2016, le projet MesInfos se fixe pour objectif de constituer à la fois les bases (techniques, économiques et juridiques) et une dynamique d'innovation autonome autour de ce que l'on désigne aujourd'hui comme "Self Data" ou encore "Personal Big Data". Il s'agit bel et bien de contribuer à l'émergence d'un marché entièrement nouveau, celui de la production, l'exploitation et le partage de données personnelles par les individus, sous leur contrôle et à leurs propres fins.

Cette action devra travailler sur :

- **Des briques techniques et économiques**, à commencer par des architectures décentralisées d'administration des données personnelles par les individus ("cloud personnel" et/ou serveur personnel "de confiance" pour du Big Data en "*peer to peer*"), permettant l'émergence autour d'elle d'un écosystème d'applications et de services tiers.
- **L'émergence d'un "espace de confiance"** capable d'assurer, en particulier, l'identification des utilisateurs, la sécurisation des échanges, mais également la portabilité des données, l'interopérabilité des plates-formes, un rôle de "place de marché" pour les applications assurant aussi un certain niveau de supervision (traçabilité des données).
- **Le soutien à l'émergence et l'expérimentation d'applications**. A titre d'exemple, le domaine de la banque et de l'assurance pourrait constituer un terrain fécond pour démarrer. Le projet MesInfos a ainsi, par exemple, identifié le potentiel de relevés bancaires consolidés (interbancaires), "augmentés" (chaque écriture donne accès à des éléments d'information sur ce qui justifie son existence : contrat, facture ou ticket de caisse, emails, déplacements, entrées d'agenda...) et restituées de manière synthétique (analyse budgétaire, conseils...). D'autres points de départ ont également été identifiés dans des domaines tels que la santé et le bien-être (en lien avec le développement du *quantified self* et dans la lignée du *Blue Button*), l'énergie, la mobilité, la gestion du quotidien ("coffres-forts numériques"), le commerce (VRM).

En outre, le projet MesInfos se fixera pour objectifs :

- D'aider les entreprises et acteurs publics à proposer des "challenges" à des entrepreneurs, chercheurs, développeurs, afin d'inviter les innovateurs à se concentrer sur des problèmes "réels", identifiés par des acteurs.
- D'explorer des nouvelles pistes d'application au travers d'ateliers, monographies, mémoires et recherches.
- D'identifier et de qualifier les défis juridiques, économiques et techniques qui conditionneront le développement à grande échelle du Personal Big Data : verrous technologiques, standards, modèles économiques, interprétation ou évolution des règles, etc. Les réponses à ces défis seront généralement européennes, voire mondiales. MesInfos poursuivra donc sa collaboration avec les innovateurs du Self Data, notamment aux Etats-Unis et au Royaume-Uni. Cet effort pourra également déboucher sur la soumission de projets européens de R&D ou d'innovation.

Action	Projet MesInfos
Pilote	Fondation Internet Nouvelle Génération (FING), en collaboration avec CozyCloud
Acteurs concernés privés	AXA, La Banque Postale, Crédit Coopératif, CozyCloud, Intermarché/les Mousquetaires, Orange, Société Générale, Solocal Group
Levier(s) public(s) à actionner	-
Acteurs concernés publics	-
Calendrier de mise en œuvre	Lancement du projet Septembre 2014 Livrables : - Architectures décentralisées et sécurisées d'administration des données personnelles par et pour les individus, sous leur contrôle et à leurs propres fins. - Expérimentation d'applications mettant le potentiel du Big Data à portée des individus.

Action 5.3 : Création d'une plate-forme ouverte de développement de services autour des objets connectés : Datavenue

Contexte

Suivant de près l'émergence rapide des smartphones, l'Internet des Objets connaît, depuis la mi-2013, une très forte accélération de son développement. L'apparition de nombreux objets connectés intégrant des capteurs très variés va entraîner de profondes évolutions dans de nombreux domaines en particulier celui de la santé, du bien-être et de la forme physique, de la voiture, du domicile, de la sécurité.

Cet essor va s'accompagner de l'apparition de nouveaux services et d'applications et va également amplifier la prolifération des données. Certaines seront des données personnelles, ce qui implique de s'assurer que ces nouvelles activités seront mises en œuvre en appliquant strictement les différentes exigences relatives au respect de la législation informatique et liberté (information et consentement de l'utilisateur, droit d'accès et droit à l'oubli, sécurité, anonymisation le cas échéant).

Plus précisément, la multiplication des objets connectés va accélérer l'émergence de gisements de données pour de nombreux acteurs économiques (opérateurs télécom, banques, assurances, industriels, distributeurs, transporteurs..) dans chacun des domaines.

La rencontre des données issues de ces objets connectés, qu'elles proviennent, de grands groupes ou d'autres acteurs, peut permettre de constituer des collections inédites de type Big Data, dont le volume, la précision, la richesse et la portée seront à la source d'énormément d'analyses poussées, d'opportunités de croisement et de corrélations par l'intermédiaire de services et d'applications qui sauront « révéler » des informations de plus haut niveau.

Quelques facteurs clés de succès

Le développement des usages des objets connectés sera intimement lié à la richesse des services proposés autour de ces objets ainsi qu'à la simplicité d'accès à ces services.

En particulier, face à la multiplication des objets connectés, chaque consommateur sera demandeur d'applicatifs permettant à la fois de :

- piloter tous leurs objets au travers d'une interface unifiée ;
- stocker et sécuriser leurs données ;
- gérer et contrôler leurs données et leurs profils ;
- augmenter l'usage et la durée de vie des objets par l'intermédiaire de services donnant une pertinence quotidienne à chaque objet.

Pour construire un écosystème viable autour des objets connectés, la participation d'une masse critique de partenaires (grands acteurs économiques et start-up) sera indispensable alors que des acteurs internationaux (éditeurs de logiciels, constructeurs et équipementiers, grands de l'internet...) commencent à déployer leur propre écosystème. Un déploiement rapide et agile est par conséquent une condition du succès.

Projet de plate-forme ouverte pour le développement de services autour des objets connectés

Porté par Orange, le projet Datavenue vise à constituer une plateforme agrégeant les données issues des objets connectés, qui sera ouverte à de nombreux développeurs tiers dans une démarche d'open innovation, pour proposer les services de demain, mashups de données et objets connectés.

Ceci implique la mise en place d'une infrastructure de données pour collecter, stocker, agréger et sécuriser les données provenant des grands acteurs économiques. La construction de l'écosystème autour de cette plateforme se fera avec l'aide des grands groupes et des fabricants d'objets connectés - dont beaucoup des plus innovants sont français - pour accueillir des milliers de développeurs tiers de services de demain.

Les données seront accessibles au travers d'API à l'attention de développeurs tiers qui pourront concevoir des mashups (programmation informatique rapide par combinaison d'APIs) prenant la forme de services en ligne ou d'applications. Cette architecture permettra d'assurer que l'exposition des données se fera selon les règles fixées par les entreprises qui les fournissent tout en respectant les contraintes relatives à la législation sur la protection des données personnelles.

Bien entendu, la quantité d'informations et le caractère personnel de la majorité d'entre elles rendront tout à fait cruciale la question du stockage et des modalités d'utilisation de ces données : il sera essentiel non seulement de garantir la sécurité des données collectées ou produites, mais aussi la transparence de leur gestion et de s'assurer que celles-ci restent à tout moment sous le contrôle des clients.

Pour ce faire, les données seront stockées en France.

Exemples potentiels d'application

1. Mes Données : tableau de bord personnel qui permettrait à un utilisateur de visualiser l'ensemble des données qui le concernent et de rechercher des corrélations entre des domaines qui ne sont pas forcément liés (par exemple l'activité internet, le sommeil et l'environnement, ...).
2. Bien être en direct : base de référence collectant anonymement des données physiologiques et d'activité par catégorie de population et par entité géographique permettant de savoir détecter l'évolution de facteurs de risques en temps réel (sommeil, stress, hyperconnectivité, problématiques d'attention, conduite dangereuse, ...) à l'attention de tous.

Problématiques soulevées

Ce projet soulève l'ensemble des problématiques déjà évoquées dans le présent document et inhérentes au Big Data :

- respect de la vie privée notamment au travers du consentement des clients et de l'utilisation de modèles d'anonymisation ;

- sécurité des données et des objets ;
- interopérabilité entre les objets et les plateformes, avec en filigrane la création de standards de collecte au niveau national et européen ;
- identification et mise en œuvre de modèles économiques novateurs ;
- facilitation de l'accès aux données des grands acteurs (open data et big data).

Bénéfices attendus

Ce projet devrait permettre :

- pour les utilisateurs d'objets connectés, de bénéficier d'une offre de service plus riche et respectueuse de la vie privée de chacun ;
- pour les start-up, d'accéder à des gisements de données permettant de développer leurs offres ;
- pour les acteurs industriels français participant au projet, de consolider leur position sur un marché agressif tout en valorisant leur savoir-faire.

Calendrier et avancement

Le programme prévoit sur 2014/2015 le développement de l'infrastructure, des services, des outils digitaux et la communication nécessaires. La démarche de création d'un écosystème devra être soutenue par l'État afin de fédérer les acteurs économiques, de promouvoir la normalisation, de stimuler le développement des services et de garantir la pertinence des engagements sur le stockage et le traitement des données personnelles.

En particulier, Orange conduit depuis le printemps 2014 des échanges avec un certain nombre de ces acteurs sous la forme d'ateliers et a reçu à date l'accord de principe de plusieurs grands comptes et startups fournisseurs d'objets et de services connectés. Une démarche « ouverte » a en outre été initiée pour la mise en place de quelques facteurs de succès de l'écosystème : normalisation de l'internet des objets, transparence vis-à-vis des utilisateurs et respect de la vie privée, règles de sécurité pour l'accès aux données, ouverture vers les développeurs au travers d'APIs.

Un challenge a été lancé début octobre 2014 par Orange et ses partenaires. Ouvert au public, son objectif était d'identifier des cas d'usage et des bonnes idées mettant en jeu des objets connectés. 320 idées ont ainsi proposées par les participants spontanés à ce challenge, au travers de la plateforme d'open innovation « Imagine with Orange (imagine.orange.com).

À partir de ces idées, une centaine d'entrepreneurs ont proposé des projets de nouveaux services ou objets connectés, dont les plus prometteurs seront récompensés le 4 décembre prochain par les partenaires de Datavenue.

Les entreprises lauréates seront ensuite accompagnées pour concrétiser leurs projets, en utilisant la plateforme Datavenue avec un objectif de réalisation complète d'ici la fin du semestre 2015.

Si les résultats de cette première étape sont à la hauteur des espoirs, l'objectif est ensuite de constituer un programme partenaires pérenne, au cours du premier semestre 2015, afin d'accueillir de nouvelles entreprises et d'ouvrir largement des interfaces d'accès aux données (API).

Action	Création d'une plate-forme ouverte de développement de services autour des objets connectés
Pilote	Orange
Acteurs privés concernés	Aleph one, Altran, Malakoff-Médéric, Mars Petcare, Netatmo, Orange, Schneider Electric, Seb, Société Générale, Suez Environnement, TF1
Levier(s) public(s) à actionner	-
Acteurs publics concernés	DGE
Calendrier de mise en œuvre	Lancement du projet Septembre 2014

Action 5.4 : Smart City energy analytics

CONTEXTE :

La gestion de l'énergie est une préoccupation majeure pour les villes, notamment pour des raisons environnementales. Avec la transition énergétique et son objectif d'intégrer progressivement dans le mix total de nouvelles sources d'énergie renouvelable, la production d'énergie, le transport et les réseaux de distribution, connus sous le nom « grids », vont évoluer d'une architecture verticale en régime prévisible à une architecture de plus en plus horizontale en mailles, « peer- to peer », et de moins en moins prévisible. Il est donc primordial pour les villes d'adopter une gestion plus intelligente afin de gérer au mieux sa consommation mais également sa production d'énergie.

Dans une telle évolution, la gestion des données devient un enjeu majeur pour les opérateurs d'électricité, afin de tirer parti de la croissance exponentielle des données produites par les différentes parties prenantes liées à l'énergie : les données sur la consommation individuelle d'énergie; sur la mobilité des personnes ; des véhicules électriques ; de disponibilité et de production des sources d'énergie renouvelables ; etc.

Grâce aux nouvelles technologies « Big Data », et leur capacité à extraire du sens des comportements ou des capacités globales et locales, il deviendra possible d'obtenir une interaction entre les bâtiments, les quartiers, les transports publics et les véhicules électriques en tenant compte des contraintes du réseau électrique.

Les acteurs de la ville seront devenus plus « intelligents » en raison de leur capacité à mesurer et agir de manière plus pertinente.

Objectifs du projet de recherche, défis et débouchés:

Le projet SCE cherche à développer une plate-forme ouverte d'analyse de données associant les fournisseurs de technologies, les intégrateurs de systèmes, les services énergétiques et de transport, les opérateurs et les entités de recherche universitaires. Cette plateforme permettra ainsi de tester différentes stratégies de gestion énergétique et éventuellement de voir apparaître de nouveaux business modèles.

Le projet se décline en quatre axes majeurs :

- Collecte/Analyse/Amélioration de la qualité de données hétérogènes (capteurs, réseaux sociaux, etc.)
- Modélisation du comportement énergétique des systèmes mis en jeu
- Aide à la décision au travers d'outils de restitution/visualisation innovants, fiables et pertinents
- Développement de business modèles

Les défis du projet consistent à:

- Définir et choisir les architectures et les technologies de collecte et d'analyse de données

- Rechercher de nouvelles stratégies en matière de visualisation de données pour l'aide à la décision.
- Évaluer le potentiel de briques technologiques développées sur différents cas d'utilisation et les intégrer.
- Assurer l'interopérabilité sécurisée avec d'autres infrastructures de la ville (transports, éclairage, bâtiments)

Les débouchés de ce projet concernent le secteur de l'énergie et plus spécialement :

- les opérateurs du réseau électrique (producteurs, distributeurs)
- les pouvoirs publics (gestionnaire de ville, etc.)
- les « prosommateurs »
- les opérateurs de transport

Support financier de partenaires industriels:

Alstom Grid (grand groupe) <http://www.alstom.com/grid/fr/>
 Alstom Transport (grand groupe) <http://www.alstom.com/transport/fr/>
 GDF SUEZ (grand groupe) <http://www.gdfsuez.com/groupe/en-bref/>
 OVH (PME Française) <http://www.ovh.com/fr/index.xml>
 Sherpa Engineering (PME Française) <http://www.sherpa-eng.com/>
 The CosMo Company (PME Française) <http://www.thecosmocompany.com/>
 Artelys (PME Française) <http://www.artelys.com/fr/accueil>
 Ecogelec (PME Française) <http://www.qualit-enr.org/ecogelec>
 G2 Mobility (PME Française) <http://www.g2mobility.com/>

Support scientifique de partenaires académiques:

CEA List
 CentraleSupélec
 Ifsttar

Organisation des activités en tâches :

Les activités du projet sont organisées en 6 tâches:

- Etat de l'art
- Data management
- Modélisation & Simulation
- Analytics & Visualisation
- Plateforme
- Valorisation & Modèles économiques

Dimensionnement et durée du projet :

Le projet aura une durée de 4 ans, pour un effort total d'environ 48 hommes-an.

Action	Projet de recherche Smart City Energy analytics
Pilote	IRT SystemX
Acteurs privés concernés	IRT SystemX, GDF SUEZ, Alstom Grid, Alstom Transport, Sherpa Engineering, The CosMo company, Ecogelec, G2 Mobility, OVH, Artelys
Levier(s) public(s) à actionner	-
Acteurs publics concernés	CEA, Centrale / Supélec, Ifsttar
Calendrier de mise en œuvre	Lancement du projet effectué le 17 Septembre 2014

Action 6 : Projets sectoriels dans le secteur public :

Action 6.1 : Secteur de l'Emploi

Objectifs et motivations

Le secteur de l'emploi est un **champ d'application particulièrement adapté pour le big data** :

- La baisse du chômage est la principale priorité du gouvernement ;
- En tant que responsable des politiques de l'emploi et de formation professionnelle, l'Etat est totalement légitime pour impulser une initiative « Big Data » dans ce domaine ;
- Il existe de très nombreuses données disponibles sur les actifs en emploi ou en recherche d'un emploi et sur les entreprises, qui pourraient être davantage exploitées ;
- Le développement du numérique fait partie des priorités de Pôle emploi qui a pris de nombreuses initiatives récentes dans ce domaine et souhaite poursuivre dans cette voie : création d'un accompagnement 100% Web pour certains demandeurs d'emploi, mise en place d'un agrégateur d'offres d'emploi avec des acteurs de l'emploi en ligne partenaires (projet de « transparence du marché du travail »), etc.

Action proposée

Premier projet lancé à court terme : Mise en place de Challenges Big Data permettant d'alimenter l'Emploi Store² avec des applications mobiles, véritables outils d'aide à la décision pour le demandeur d'emploi dans sa recherche d'emploi et pour le recruteur dans sa recherche de candidats

L'Emploi Store[®] constitue un levier pour favoriser l'émergence du sujet Big Data. Le projet a deux composantes sur lesquelles pourraient s'appuyer des projets Big Data :

- un dispositif d'ouverture et de partage des données de Pôle emploi à des acteurs externes, via une plateforme dédiée à la mise à disposition de données
- le recours à des « sur-traitants » permettant de gagner en agilité et de s'inscrire dans une logique d'innovation collaborative

Pour accélérer la démarche Big Data, Pôle emploi a placé celle-ci dans le cadre de l'Appel à Manifestation d'Intérêt « Cœur de filière numérique – challenges Big Data », lancé le 12 février 2014.

Ainsi, 2 challenges vont être lancés avec CapDigital en 2015 : l'objectif est d'organiser des hackathons afin de mettre à disposition des données à des développeurs, favorisant le développement d'applications autour de thématiques « emploi » à définir, qui pourraient être référencées dans l'Emploi Store[®].

² L'Emploi Store[®] est un projet qui vise à concevoir et développer une plateforme de mise à disposition de données, qui permettra à Pôle emploi et aux communautés de développeurs de proposer de nombreuses applications pertinentes sur la thématique de l'emploi aux demandeurs d'emploi, actifs et recruteurs.

Action	Projets dans le domaine de l'emploi
Pilote	Pôle Emploi
Acteurs privés concernés	Les contributions d'entreprises tierces dans le cadre des challenges restent à identifier
Levier(s) public(s) à actionner	Appel à Manifestation d'Intérêt « Cœur de filière numérique – challenges Big Data »
Acteurs publics concernés	Le pilotage et le sponsoring seront assurés par Pôle emploi. BPI
Calendrier de mise en œuvre	L'action sera mise en œuvre opérationnellement en 2015 avec 2 challenges conçus au premier trimestre et effectifs au deuxième trimestre.

Action 6.2 : Secteur de la Santé - Programme de sensibilisation, de repérage et d'accompagnement des assurés MGEN portant sur la sphère cardio-vasculaire

Objectifs et motivation

Au regard des enjeux actuels du système de santé (soutenabilité financière, chronicisation des pathologies, vieillissement de la population) la MGEN souhaite s'engager en coordination avec les pouvoirs publics, dans la définition et la mise en œuvre d'une stratégie de gestion du risque santé.

Cette stratégie permettra notamment au Groupe de proposer à ses assurés une offre structurée et articulée de services en santé tenant compte des spécificités de la population couverte au travers de programmes de prévention globaux des principales maladies chroniques. La démarche de la MGEN a également pour ambition l'efficacité des parcours de santé et la maîtrise de la sinistralité RO/RC. A ce titre, elle s'inscrit en complément de celle de la CNAMTS qui a défini une politique ambitieuse de gestion du risque santé.

Pour la MGEN, cette stratégie participe des missions qui lui sont confiées dans le cadre de la délégation de gestion du régime obligatoire d'assurance maladie.

Un premier programme de sensibilisation, de repérage et d'accompagnement est proposé en ce sens. Actuellement en construction, son déploiement est prévu pour le premier semestre 2015. Il concerne la sphère cardio-vasculaire incluant ses facteurs de risque, tant métaboliques que liés aux comportements de vie de la personne (alimentation, tabac, activité physique), et ses principales pathologies (hypertension, diabète, maladie coronarienne, insuffisance cardiaque, accident vasculaire cérébral).

Il permet d'accompagner l'assuré dans le cadre d'un parcours de santé individualisé grâce à 4 étapes d'intervention :

- l'information et la sensibilisation ;
- le repérage des facteurs de risque cardiovasculaires ;
- l'accompagnement à la maîtrise des facteurs de risque : e-learning, e-coaching, télécoaching,, télésurveillance des données de santé au domicile ;
- le suivi du plan thérapeutique et de l'observance.

Le Programme génère certes un investissement humain, structurel et financier important mais également un retour sur investissement positif du point de vue médico-économique (économies de coûts de santé, économies de coûts de prévoyance)

Le Programme a vocation à être généralisé et donc industrialisé dans son déploiement : celui-ci est prévu dans un 1^{er} temps sur 14 régions en 5 ans, après une phase test de 2 ans sur deux régions pilotes et une évaluation clinique, médico-économique, socio-organisationnelle et de service

médical rendu.. Il intégrerait en première estimation 20 000 assurés en moyenne en mode pilote, sous l'égide de 2 ARS.

Le concept

La MGEN, en tant que Mutuelle professionnelle, référencée par le Ministère de l'Education Nationale, souhaite ainsi proposer aux pouvoirs publics de déployer un premier programme d'accompagnement des patients. Ce Programme se caractérise par :

- une base populationnelle circonscrite et bien identifiée permettant le déploiement rapide de différents modules : la sensibilisation des assurés aux problématiques de santé cardiovasculaire, l'évaluation du risque cardiovasculaire global, le télésuivi des assurés au domicile par télé-médecine en relation avec le médecin traitant et un accompagnement individualisé dispensé par des professionnels de santé d'une plateforme médicalisée ;
- le traitement de données de remboursement permettant une segmentation fine de la population et des risques, et la construction d'un modèle médico-économique probant ;
- l'organisation d'un parcours de santé sur l'ensemble de la sphère cardio-vasculaire renforçant significativement la prévention des risques ;
- la valorisation du rôle du médecin traitant et des prises en charge par le 1^{er} recours ;
- une vocation nationale sur la base d'un déploiement progressif et de temps d'évaluation partagés.

Le Programme est conçu pour consolider la relation patient-médecin et non la remplacer. Le médecin traitant de l'assuré est ainsi prescripteur et acteur du programme :

- Il propose et est le seul à pouvoir entériner l'adhésion de l'assuré au programme
- Il a la maîtrise du contenu du plan thérapeutique et de son suivi par son patient : il peut le consulter à tout moment, notamment au cours de la consultation
- Il est informé en continu du déroulement du programme, du suivi de son patient et de l'évolution de ses comportements (hygiéno-diététiques ou observance).
- Il peut être, à sa demande, destinataire alertes de suivi des constantes médicales à distance de son patient.

Le Programme participe également *de la transition de la posture de patient « Client » ou « Usager » à celle de patient « Acteur » de sa santé* en renforçant leur communication avec leur entourage médical, en leur apportant une meilleure compréhension de leur maladie et de sa prise en charge, en suivant leurs données de santé et en spécifiant leur accompagnement sur la base d'un questionnaire motivationnel. Il permet également au médecin d'utiliser au mieux le temps de la consultation en orientant son patient vers des modules d'accompagnement adaptés à sa prise en charge.



Bénéfices attendus

Le Programme de gestion du risque cardio-vasculaire de la MGEN présente des bénéfices multiples pour le Groupe mais également pour l'économie française au sens large :

- **Bénéfice sociétal :** amélioration de la santé de sa population cible Education Nationale, par le biais de programme de gestion du capital santé
- **Bénéfice économique :** impact positif sur le coût de la protection sociale en France (santé, arrêt de travail) par évitement de dépense
- **Création d'emploi :**
 - **Secteur médical :** développement de nouveaux métiers à forte valeur ajouté au travers des plateformes médicalisées : téléopérateurs spécialisés en télésanté, infirmières, diététiciennes, éducateurs médico-sportifs formés à la e-prévention et au coaching à distance, sous la responsabilité de médecins de plateforme. Ce type d'activité de plateforme technologique est à très forte valeur ajoutée et haut niveau d'expertise et donc non délocalisable. L'émergence de ces métiers pourrait ouvrir de nouvelles voies d'évolution des parcours universitaires des facultés de médecine.
 - Ajouter les professionnels de l'ingénierie e-santé pour piloter ces programmes, les déployer, les animer, les coordonner ???
 - **Secteur Statistique:** La dimension statistique des moteurs de génération d'algorithmes liés aux facteurs de risque pourrait conduire à l'émergence de nouveaux emplois dans le domaine de la statistique mathématique et le développement de filières d'excellence de statistique et d'économétrie dans la cadre des filières d'actuariat (du type ENSAE Data Scientist).
- **Innovation territoire de santé numérique:** les volets « quantified self » et télémédecine du Programme encourageront le développement de solutions innovantes de monitoring à distance de la santé (tracker d'activité, auto-tensiomètre communicant, balance communicante...)

Ce Programme pourrait s'inscrire dans une vision plus globale de standards technologiques (du type Continua Alliance) et médicaux pour le développement de services en santé dans le cadre Européen et dans le contexte de programme du type United4Health par exemple.

D'autres opérateurs du Régime Obligatoire se sont déjà rapprochés de la MGEN et ont exprimé leur intérêt à rejoindre le programme et à mutualiser les services. Des groupes de travail sont en cours.

Pilotage et contributions

Attente d'engagement des partenaires

Le Programme de gestion du risque cardiovasculaire de la MGEN sera proposé gratuitement aux professionnels de l'éducation nationale français volontaires dans les zones pilotes identifiées.

Le coût de construction du programme inclus notamment :

- Développement de la plateforme numérique et du système d'information (site internet, dossier patient, moteur de règles des outils d'accompagnement, dispositifs communicants,...)
- Création de la plateforme téléphonique médicalisée (infrastructure de téléphonie, recrutement des professionnels de santé, formations, ...)
- Développement des spécifications médicales, des processus d'inclusion des assurés, de l'approche comportementale d'accompagnement et des contenus de coaching
- Analyse médico-économique de validation du modèle d'accompagnement
- Développement du modèle statistique de stratification de risque
- Développement des outils de communication adaptés

Attentes d'engagement de l'Etat

L'efficacité du Programme d'accompagnement de la MGEN réside dans sa capacité à recourir à un algorithme, capable de traiter les données de liquidation des prestations en santé³, pour prédire des états de santé cardio-vasculaire des assurés. Cet algorithme est aujourd'hui opérationnel et fiable à 85%.

Pour ce faire, la MGEN doit signer avec la CNAMTS une convention spécifique pour conduire l'expérimentation du programme dans le cadre de l'article L162-1-11 du Code de la sécurité sociale et relever du décret N°2012-1249 du 9 novembre 2012 relatif aux traitements des données à caractère personnel pour la mise en œuvre des programmes de prévention et d'accompagnement en santé des assurés sociaux.

Aujourd'hui, seule la CNAMTS relève de ce décret pour deux de ses programmes d'accompagnement : généralisation du programme Sophia et déploiement du programme d'accompagnement du retour à domicile des patients hospitalisés (PRADO).

Il est attendu une clarification de la part des autorités compétentes, permettant à la MGEN d'utiliser ces données dans le cadre du projet présenté.

La MGEN, au titre de sa délégation de gestion du Régime Obligatoire, a adressé à la CNAMTS le 30 septembre 2014, un projet de convention spécifique. Dans l'attente d'un retour favorable, la MGEN propose un mode d'inclusion dégradé moins efficace et plus coûteux, comme le démontrent les études universitaires en la matière dans les pays anglo-saxons. Il repose sur la

³ La MGEN dispose des données de liquidation dans des bases de données sécurisées et strictement confidentielles.

seule utilisation des données de questionnaires d'évaluation du risque cardiovasculaire global ou de campagnes de repérage sur site, pour inclure les assurés se portant volontaires pour s'engager dans le Programme. Par ailleurs, le recrutement volontaire des patients dans un programme d'accompagnement amène un biais d'évaluation.

Action	Programme de sensibilisation, de repérage et d'accompagnement des assurés MGEN portant sur la sphère cardio-vasculaire
Pilote	MGEN
Acteurs concernés privés	Partenaires opérationnels sélectionnés
Levier(s) public(s) à actionner	Réglementaire (utilisation des données assurés santé)
Acteurs concernés publics	MGEN
Calendrier de mise en œuvre	Mai 2014 : Clarification des aspects réglementaires Septembre 2014 : Identification des deux régions pilotes Septembre 2015 : premiers indicateurs d'inclusion des assurés dans le Programme et lancement des séances de coaching Juin 2016 : Rapport intermédiaire d'évaluation Septembre 2016 : Préparation de la généralisation

Action 6.3 : Filière génomique - Séquençage génomique et oncologie

Objectifs et motivations

Les techniques de séquençage à haut débit ou séquençage massivement parallèle sont apparues en 2006 et sont désormais utilisées de façon standardisée en médecine. Le coût/délai du séquençage d'un génome est passé de cinq ans et trois milliards de \$ (2000) à dix jours et dix mille euros en 2012, avec des perspectives de génome à 1000 euros désormais envisageables.

Cette démocratisation du séquençage offre désormais des perspectives d'utilisation de ces technologies à large échelle et à relativement court terme. Le Plan Cancer 3 présenté en avril 2014 en France fait par exemple état de 50 000 séquençages réalisés sur des patients à l'horizon 2019, et prévoit le lancement d'initiatives d'expérimentation dès 2014. En revanche, la volumétrie correspondant à ces données est significative : un génome complet correspond à 200 giga octets, il convient de séquencer plusieurs génomes par patient (tissu sain et tissu tumoral en cancérologie par exemple), et séquençer 50 000 patients revient donc à devoir gérer de l'ordre de 10 à 20 peta octets... sans compter toutes les autres possibilités en dehors de la cancérologie

La génération de ces données, et l'intérêt clinique qu'elles suscitent pose désormais trois questions principales : 1) Comment gérer ces volumes de données; 2) Comment tirer parti de ces données et permettre au clinicien de les interpréter et 3) comment mettre en place des filières de séquençage sur le territoire

- 1) Big Data : La gestion des volumes de données demeure un challenge significatif, non seulement du fait des volumétries, mais également de la nature des informations médicales, par définition sensibles, propriété du patient, et devant être particulièrement protégées -
- 2) L'interprétation des données de génomique fait l'objet de nombreuses initiatives, en particulier aux USA (Caris, Fondation Medicine, ...) ou Européennes (Qiagen/ingenuity) mais également en France, où un consortium regroupant Gustave Roussy, l'Inserm, IntegraGen et Sogeti HiTech a lancé le projet ICE (interpretation of clinical exome) afin de développer un outil logiciel d'assistance à l'interprétation de ces données. Ce projet a été soutenu par l'Etat dans le cadre de l'appel à projets FUI 17
- 3) La troisième dimension, celle de la filière industrielle du séquençage doit également être prise en compte, de façon à pouvoir établir les plateformes intégrées de séquençage, en favorisant le transfert des entités de séquençage du domaine hospitalier à celui de l'industrie, seul capable de gérer la complexité de la filière (gestion des échantillons, opérations de séquençage, génération des données en qualité et en délai, formatage des données utilisables). A l'initiative du Génopole d'Evry, un projet constitué de Génopole, IntegraGen, Sanofi, le CEA, Sogeti HiTech et de plusieurs entités cliniques (ICAN, IHU, Mondor notamment) a répondu à au concours mondial innovation avec un projet destiné à structurer les filières de séquençage au niveau national.

A ce jour, l'aspect Big Data / gestion des données incluant la définition des infrastructures nécessaires, des processus de gestion, d'accès, de conservation et d'utilisation des données ne fait

pas l'objet d'une initiative structurante, ce qui pourtant devrait être une des composantes majeures de la mise en pratique de la génomique sur le territoire.

Bénéfices attendus

Les bénéfices attendus sont de natures sociétales et économiques. L'interprétation des données génomiques par les cliniciens permettrait d'éviter de l'ordre de 20% de traitements lourds (type chimiothérapie) inadaptés aux patients diagnostiqués.

Ces patients ne subiront plus les effets secondaires indésirables de traitements par ailleurs sans effet sur leur pathologie au regard de leurs spécificités génomiques.

En termes économiques, chaque traitement inadapté, évité, permettrait d'économiser de l'ordre de 30 à 50 000 euros. Ces montants pourraient être réutilisés dans de nouveaux traitements, mieux adaptés aux particularités de chaque patient.

Pilotage et contributions

Les projets visés par cette action (autre que le projet ICE) seront conduits en collaboration avec le plan Hôpital numérique.

Action	Génération et interprétation des données de génomique
Pilote	Integragen
Acteurs privés concernés	Integragen, Sogeti Hi Tech
Levier(s) public(s) à actionner	Financement dans le cadre de l'appel FUI 17
Acteurs publics concernés	Institut Gustave Roussy, INSERM
Calendrier de mise en œuvre	Mai 2014 : Lancement officiel du projet Jalon 1– Q4 2014: Release 1 / Minimally viable product Jalon 2 (Q3 2015) : V1 du logiciel avec retour cliniciens Jalon 3 (2016) : Commercialisation

Actions de nature à modifier la réglementation autour du Big Data en France

Action 7 : Evolution du contexte réglementaire et législatif

Objectifs et motivations

Le Big Data est souvent décrit comme l'ensemble des technologies permettant la valorisation des données massives, mais c'est aussi, et surtout, une arme de guerre des grandes sociétés du numérique pour s'attaquer méthodiquement à des pans entiers de l'économie (industrie et services) qui n'ont pas encore été impactés par la révolution numérique. A l'instar de ce qui s'est passé pour la presse, le commerce de proximité, le tourisme ou encore l'industrie musicale, ces sociétés tentent actuellement d'utiliser 1/ les centaines de millions de clients avec lesquels elles ont une relation de proximité et de confiance, 2/ les nombreuses données qu'elles ont sur ces clients, et 3/ leur capacité de traitement et de croisement de ces données pour modifier profondément le paysage concurrentiel, et en particulier la relation client et la chaîne de valeur. C'est ainsi que des secteurs comme l'assurance, la banque, la grande distribution, le crédit à la consommation, l'industrie automobile, les « utilities », l'énergie, et bien d'autres encore risquent de voir leur positionnement dans la chaîne de valeur réduite progressivement au rôle de sous-traitant « technique », avec une très forte pression sur les marges et un risque potentiellement létal pour les plus faibles.

L'enjeu disruptif potentiel sur l'économie française et un grand nombre de ses industries et services est donc tout à fait considérable. Et le Big Data reposant sur la réutilisation à des fins différentes et le croisement de données existantes ou nouvelles, il résulte que le développement du Big Data 1/ va concerner un nombre croissant de sociétés et 2/ va engendrer des usages de plus en plus complexes et sophistiqués des données au sein de « processus industriels complexes » impliquant potentiellement une grande diversité d'acteurs.

Afin d'encourager à la fois 1/ une simplification des formalités préalables auprès de la CNIL dans une optique de mise en conformité, et 2/ un renforcement de la sécurité juridique pour les acteurs impliqués dans des processus industriels complexes, sur les questions informatiques et libertés, il est proposé de s'engager dans la voie de la certification de processus industriels complets liés à des scénarios détaillés d'usage et de croisement de données. Ces « *certificats de conformité de processus* » seraient donc en quelque sorte une forme de « rescrit juridique » pour des usages bien identifiés qui pourraient être utilisés par les industriels dans une logique de « label éthique » leur permettant de promouvoir leurs usages éthiques auprès des utilisateurs.

Action proposée

L'approche proposée consiste à définir de nouveaux « Packs de conformité », déjà testé par la CNIL, et à les renforcer par un processus de certification de processus industriels. Une des premières étapes sera de publier les règles qui permettent de certifier un processus industriel.

Cela consiste à promouvoir une approche « métier », par industrie, et surtout par type d'usage à l'intérieur de chaque « écosystème industriel » né de ce nouvel usage, comme par exemple l'écosystème autour de la voiture connectée (constructeurs automobiles, assureurs, systèmes de

navigation, etc.). A chaque « processus industriel » serait associé une sorte de « norme » décrivant en détail les données utilisées, le croisement qui en est fait, les usages qui en sont dérivés, les calculs effectués, les processus d'anonymisation employés, etc. Ces normes pourraient être proposées par les industriels eux-mêmes, ou par les organisations professionnelles, et seraient étudiées, discutées, et bien entendu validées par la CNIL conformément aux dispositions de l'article 11.3° a) de la loi. La CNIL pourrait ensuite, si elle le juge opportun, adopter des dispenses de formalités associées à ces normes dans le cadre des « packs de conformité ». L'adhésion à ces normes serait volontaire, et en cas de non adhésion à l'une des normes existante, les entreprises seraient bien entendu toujours soumises au droit commun par l'application de la loi au cas par cas.

L'un des objectifs additionnels de cette approche est de faire des normes de véritables « labels éthiques », et promus en tant que tels par chaque industrie, pour mettre en avant le caractère éthique de leurs produits. Cette mise en avant du caractère respectueux de la vie privée devra être un des atouts essentiels des sociétés européennes pour se positionner face aux géants du numériques, notamment américains, qui par le caractère multiple de leurs activités, et le croisement des données associées, auront beaucoup plus de mal à adopter une telle approche éthique.

L'avantage de cette approche par « normes » est qu'elle pourrait aussi permettre, à terme et d'une certaine manière, à la CNIL de « certifier les certificateurs » (des sociétés comme VERITAS par exemple), dès lors que ceux-ci vérifieraient que les conditions d'adhésion à ces normes sont bien respectées par les candidats.

A plus long terme, il est également possible d'imaginer un organisme européen (ou mondial, mais d'initiative européenne) placé sous l'égide des équivalents des CNILs des divers Etats Membres, dont le but serait de définir ces normes, et qui permettrait notamment d'exporter l'approche européenne sur le plan international, y compris aux Etats-Unis (on peut très facilement imaginer des startups américaines adoptant ces labels pour se positionner face aux géants du secteur).

Enfin, il sera peut-être nécessaire de faire évoluer la rédaction actuelle de la loi, actuellement fondée sur la notion de « finalité compatible » (afin de mieux tenir compte de la finalité du process industriel complet). Toutefois, la notion de coresponsable de traitement dans la directive européenne permet déjà largement d'appréhender cette problématique. Une autre piste intéressante pourrait consister à étudier un principe de « réciprocité » exigeant des industriels utilisant des données personnelles un bénéfice quantifiable pour les individus à l'origine de ces données, quand ces données sont utilisées ou réutilisées.

Pour finir, il ne faut pas oublier que le Big data se caractérise également par l'émergence de nouveaux services ou processus qui ne requièrent pas le traitement de données directement ou indirectement identifiantes. Ce qui compte souvent, en effet, est qu'elles soient individualisables, peu importe l'identité de la personne concernée. Afin d'assurer un développement du Big data conforme à la législation sur la protection des données personnelles, la CNIL pourrait donc « certifier » a priori des processus d'anonymisation. Les entreprises pourraient ainsi traiter les données selon un processus d'anonymisation reconnu conforme aux exigences de la loi

informatique et libertés, ce qui constituerait un atout commercial et sécuriserait le cadre juridique de leurs travaux.

Remarque : La CNIL a été consultée dans la rédaction de cette action mais n'est pas engagée par les pistes présentées.

Action	Réglementation
Pilote	CNIL
Acteurs concernés privés	-
Levier(s) public(s) à actionner	Certifications de processus industriels + Evolutions réglementaires
Acteurs concernés publics	CNIL
Calendrier de mise en œuvre	Ateliers sectoriels – Premier trimestre 2015