

# SECTION 07

## SCIENCES ET TECHNOLOGIES DE L'INFORMATION (INFORMATIQUE, AUTOMATIQUE, SIGNAL ET COMMUNICATION)

### *Président*

Bruno Durand

### *Membres*

Frédéric Bimbot

Isabelle Bloch

Ali Charara

Jean-Noël Chardron

Christine Chevallereau

Alain Denise

Étienne Dombre

Pascal Dubreuil

Chantal Enguehard

Jean-Luc Gauvain

Blaise Genest

Marc-Olivier Killijan

Bertrand Mazure

Stephan Merz

Serge Miguet

Gilles Mourot

Amedeo Napoli

Hervé Rivano

Marie Christine Rousset

Philippe Schnoebelen

Alexis Tsoukiàs

## 1. Introduction

Afin de résumer les tendances et les enjeux des sciences de l'information, nous avons adopté dans ce rapport une structure partant des thèmes de recherche fondamentaux pour aller jusqu'aux interactions avec d'autres domaines. Nous introduisons d'abord les disciplines qui soutiennent la plupart des thèmes décrits (traitement du signal, automatique, logique, sémantique, combinatoire, graphes, systèmes). Le thème du calcul effectif est ensuite développé, avec ses aspects de grilles, réseaux, ingénierie du logiciel, systèmes, vérification, sécurité et fiabilité, systèmes embarqués. Nous avons choisi deux thèmes de recherche à l'interface de l'informatique et du monde réel, avec leurs applications en ingénierie : robotique et sciences et technologies de la décision. La section suivante est dédiée aux interactions avec le monde environnant, et porte sur le traitement des images et la vision par ordinateur, les interactions homme-machine, la réalité virtuelle et augmentée, et le traitement des langues. Sont enfin développés certains thèmes interdisciplinaires dont les enjeux sont majeurs, comme la bio-informatique, le développement durable et les systèmes d'information. Nous concluons sur une brève analyse des forces et faiblesses de nos disciplines.

## 2. Fondamentaux

### 2.1 Signal

En traitement du signal, les problèmes sont de plus en plus formulés de façon intégralement numérique dans la mesure où la numérisation est de plus en plus intégrée au capteur. Ce changement de paradigme a fait évoluer le domaine de manière importante. Un des points forts des approches récentes est d'associer des modèles de représentation dans différents types d'espaces à des méthodes de modélisation et d'analyse mathématique et à des algorithmes, faisant ainsi le pont entre des aspects mathématiques et des aspects informatiques. Les orientations actuelles ont souvent des spécificités liées au domaine. Par exemple l'analyse du signal audio ou musical vise des tâches couvrant la description, la caractérisation ou la transcription automatique, la reconnaissance, la synthèse, la fouille et l'indexation, la séparation de sources, le suivi de la voix, etc.. En traitement de la parole, les méthodes d'analyse, de reconnaissance et de synthèse vont désormais de pair avec le développement du traitement des langues (voir cette partie dans le rapport), tout en gardant un ancrage fort dans la modélisation statistique. Les méthodes de traitement du signal sont également utilisées en traitement d'images, surtout pour les traitements de bas niveau tels que la restauration ou le filtrage. Cependant, les spécificités du domaine, en particulier la dimension spatiale de l'information manipulée, l'importance croissante de la prise en compte de connaissances sur le contexte, et l'évolution des problématiques vers des tâches de haut niveau (reconnaissance de structures, indexation, interprétation sémantique), ont fait de l'image et de la vision par ordinateur une discipline à part entière, et une section lui est consacrée dans ce rapport. Dans un tout autre domaine, le traitement du signal s'étend au traitement et à l'analyse de séries temporelles quelconques, avec en particulier des applications en finance. Le lien entre traitement du signal et communications numériques permet également de revisiter les questions de codage, d'estimation

et de théorie de l'information pour les nouveaux systèmes de communication. Parmi les techniques émergentes, citons les codes espace-temps, ou encore la théorie des grandes matrices aléatoires. Ce lien se retrouve dans la problématique des réseaux de capteurs, avec des questions théoriques sur le traitement statistique distribué. Le thème des représentations parcimonieuses est devenu un sujet phare, visant à faire face à la taille croissante des données en choisissant des espaces de représentation adaptés (typiquement des bases selon lesquelles le signal peut être décomposé avec un petit nombre de coefficients non nuls). Enfin, la multi-modalité se développe de plus en plus, dans de nombreux domaines, par exemple en biométrie, en fouille de données sur le web ou dans le traitement de signaux et de données dans le domaine médical.

### 2.2 Automatique

Science des systèmes, l'Automatique voit se rencontrer en son cœur les notions de système et de dynamique. Elle se fonde sur le concept d'action, dans le sens de la modification du comportement naturel d'un système, et s'appuie pour ce faire sur des thèmes fondamentaux en modélisation, identification, observation, commande et diagnostic. Elle combine des méthodologies analytiques et numériques, manipule des signaux hybrides continus et discrets, maîtrise la complexité, la dimension et l'hétérogénéité des systèmes auxquels elle se confronte. L'Automatique s'appuie sur des outils mathématiques, informatiques, de recherche opérationnelle et sur ses propres outils. Enfin, l'essence de l'Automatique réside dans la dualité entre les systèmes physiques auxquels elle s'adresse et le fait que les avancées majeures ne peuvent venir que du développement de méthodologies génériques, indépendamment de l'application pressentie. Il y aurait ainsi tout autant de risque à vouloir décliner l'Automatique sous forme d'un objet de recherche que de renoncer à la multiplicité naturelle des outils pour résoudre les problèmes que l'Automatique soulève.

Les enjeux sociétaux auxquels l'Automatique doit se confronter peuvent être ré-exprimés en termes de défis scientifiques : la maîtrise des grands systèmes dynamiques interagissant ; la prise en compte des hétérogénéités pour la commande et la prédiction ; l'assurance de la sécurité et des performances des systèmes contrôlés dans des environnements incertains ; l'intégration de nouvelles technologies dans les systèmes de contrôle commande ; la maîtrise des systèmes autonomes et embarqués, avec en particulier les aspects liés à la reconfiguration et la recalibration des lois de commande ; la réduction des coûts de validation des lois de commande. L'Automatique assume les nouveaux défis de l'information, de la communication et des nanotechnologies, à la fois au niveau de nouveaux axes de recherche associés (commande des réseaux, utilisation des nouveaux matériaux pour la commande) mais aussi d'axes applicatifs (véhicules intelligents, réseaux de capteurs, ...)

Des axes de recherche interdisciplinaire ont vocation à prendre une importance considérable dans un futur proche. C'est le cas pour les recherches en Biologie et Automatique (modélisation, observation, analyse des interconnexions dynamiques dans les systèmes biologiques). Les problèmes se déclinent dans un contexte multi-échelle, de l'ADN jusqu'au niveau macroscopique d'une population bactérienne tant dans un contexte d'analyse que de contrôle. De la biologie au biomédical, on passe de la compréhension au niveau cellulaire au développement de nouvelles technologies pour l'aide aux gestes médico-chirurgicaux, à la téléopération, au développement de prothèses et de muscles artificiels (en lien avec la robotique). En même temps les grands domaines d'interface de l'Automatique restent des moteurs puissants pour les développements méthodologiques amont, que ce soit autour de l'automobile, des véhicules aériens ou autonomes, des réseaux électriques, de l'environnement, etc.

### 2.3 Structures fondamentales

Les structures fondamentales en informatique sont les arbres, les graphes, et les tableaux. Ce sont les ingrédients de bases des algorithmes et des modèles - aussi bien d'architectures matérielles que logicielles. Différentes spécialités s'occupent de leur analyse, de leur conception, de leur utilisation ou de leur optimisation, se posant donc des questions de nature différente.

Tout d'abord, la logique s'occupe de décrire les formalismes pour exprimer des connaissances ou des propriétés de ces structures. Elle s'attelle en particulier à l'analyse de la déduction dans ces formalismes (théorie de la démonstration). Ce dernier point la rapproche du domaine connexe de la sémantique, qui abstrait la syntaxe de chaque langage de programmation en un morphisme d'une structure sémantique, unifiant de facto programmes et preuves (correspondance de Curry Howard). Une question fondamentale est de trouver les bonnes structures sémantiques, et d'analyser leurs propriétés. Un des défis de la discipline est de représenter, de manière cohérente, différents types de programmes, par exemple les programmes concurrents.

La Combinatoire énumère les objets de ces structures fondamentales ayant certaines propriétés, et tente de les dénombrer. Ainsi, il est possible de générer aléatoirement des objets, mais aussi de considérer les limites asymptotiques en moyenne d'opérateurs sur ces objets. La combinatoire s'efforce également de trouver des relations bijectives entre objets, ce qui permet en particulier de résoudre les questions d'énumération et de dénombrement, mais renseigne aussi sur la nature intrinsèque des classes de structures. Les travaux de combinatoire enrichissent depuis longtemps l'optimisation combinatoire. Elle se développe maintenant de plus en plus en liens avec la physique théorique.

La structure « graphes » joue un rôle particulier. En effet, d'une part la combinatoire des graphes est englobée dans la théorie des

graphes, qui a une très longue tradition en France, ce qui lui confère le statut de domaine à part entière. D'autre part, de très nombreux problèmes peuvent se ramener à un problème de coloration d'un graphe, les plus évidents étant des problèmes de communication. Une question fondamentale est celle de la décomposition des graphes : la représentation d'un graphe dans une structure plus simple. Les questions sous-jacentes sont de comprendre quelles classes de graphes se décomposent dans la structure la plus simple, et de calculer le coût de la transformation.

Enfin, les théories de la complexité algorithmique et de la complexité intrinsèque (Kolmogorov) sont fondamentales pour mesurer la difficulté d'un problème en termes de temps et d'espace ou de quantité d'information. Les outils utilisés sont la théorie des automates, la théorie de la calculabilité, et au delà, celle des modèles géométriques comme les pavages, modèles qui expriment la nature géométrique du calcul. Ces théories, une fois combinées à des mesures de probabilité, de temps etc., permettent aussi de spécifier ou d'analyser des phénomènes continus. Cela permet de prendre en compte l'incertitude, le comportement hybride discret/continu, de calculer des coûts en moyenne plutôt que dans le pire de cas, etc. Le mélange de ces aspects discrets et continus ouvre de nouvelles perspectives qui nécessitent des compétences variées assez rares. Une nouvelle interface entre les mathématiques appliquées et l'informatique théorique semble en émergence ici.

## **3. Réseaux, Sécurité, Algorithmique**

### **3.1 Algorithmique**

Les algorithmes sont des outils utilisés bien au-delà de la seule section 07. Ainsi, l'algorithmique a des retombées dans des domaines aussi divers que les télécommunications, l'image, la bioinformatique, la mécanique ou la physique. La plupart des algorithmes manipulent les structures

fondamentales décrites ci-dessus. D'autres portent sur des structures particulières. C'est le cas de la géométrie algorithmique, qui manipule des concepts géométriques, mais peut être utilisée dans des problèmes plus généraux. Un des nouveaux enjeux est la croissance de la taille des données à traiter et leur caractère potentiellement évolutif. Pour combattre la complexité due à la grande taille des données, des familles d'algorithmes permettent d'obtenir des solutions approchées avec des garanties de qualité. L'utilisation de méthodes aléatoires a introduit de nouvelles classes d'algorithmes dont les résultats sont « garantis » par des ratios. L'utilisation de l'aléa permet aussi de résoudre des problèmes plus rapidement que de manière déterministe. Ces algorithmes sont souvent utilisés en algorithmique du texte et du génome ainsi qu'en image, calcul symbolique ou physique théorique.

La complexité des programmes et systèmes qui traitent de la dynamique des systèmes fait de la fiabilité un enjeu difficile à accomplir. Pour cela, la vérification formelle se voit compléter par la théorie des jeux, afin de prendre en compte l'environnement, imprévisible et modélisé à des niveaux variables.

Enfin, l'algorithmique quantique produit un nouveau champ de recherche tirant partie des possibilités offertes par la physique quantique pour offrir des modèles de calculs plus efficaces. L'application phare est la décomposition en nombres premiers, qui a des implications fortes en cryptologie. Le défi est de développer une complexité quantique pour de problèmes dans d'autres champs que l'arithmétique. Si l'ordinateur quantique ne semble pas être réalisable dans un futur proche, il faut noter que l'approche quantique a permis de résoudre des problèmes algorithmiques non quantiques.

### **3.2 Réseaux**

Le caractère ubiquitaire, perversif et l'intelligence ambiante sont au cœur des réseaux de nouvelle génération. En effet, ceux-ci se diversifient et s'enrichissent en bénéficiant d'apports technologiques croisés :

réseaux de capteurs à très faible consommation, grilles de calculs à haute performance, grilles de stockage, communautés virtuelles, réseaux micro-cellulaires, etc. De cette diversité émergent des comportements et de nouveaux usages qui s'accompagnent de problématiques de passage à l'échelle en termes de complexité, dynamique et hétérogénéité, et de nouveaux besoins en termes de sécurité et de performance. Ces travaux sont l'occasion de recherches sur les grands réseaux d'interactions et des passerelles vers la physique, les sciences humaines et sociales, l'épistémologie etc.

La complexité d'Internet s'est déportée vers sa périphérie et est caractérisée par la mobilité des terminaux, la diversité de leurs capacités de communication et leur densité, rendant réaliste l'hypothèse d'utiliser ces entités terminales comme un prolongement naturel d'Internet capables de s'affranchir d'un recours systématique aux infrastructures pour l'échange d'informations entre entités terminales. Cette évolution de l'Internet des machines vers l'Internet des objets induit de nouvelles problématiques de recherche reposant sur une révolution des paradigmes de communication traditionnelle de bout en bout et de connectivité permanente. Celles-ci induisent un des nouvelles problématiques comme l'auto-organisation ou l'adressage centré sur le contenu en faisant abstraction de l'entité porteuse de l'information. Dans un tel contexte post-IP la notion d'adresse est remplacée par celle d'identificateur de donnée et la fonction de routage évolue vers le routage sémantique. De ce fait, l'optimisation et l'évaluation de performances des réseaux nécessite le développement de nouvelles approches, stochastiques ou algorithmiques, comme celles fondées sur la théorie des jeux.

Par ailleurs, les exigences sociétales imposent de prendre en compte de nouvelles métriques des réseaux pour minimiser leur consommation énergétique et maîtriser la

puissance rayonnée par les réseaux radio, notamment cellulaires. Dans le domaine des systèmes, on assiste à la convergence entre des recherches sur le parallélisme et le calcul haute performance avec les recherches sur les systèmes distribués et les intergiciels autour de la problématique de la globalisation des ressources informatiques, des infrastructures distribuées à grande échelle et de leur utilisation optimale.

### 3.3 Cloud Computing

Le cloud computing est un modèle de grille permettant à de multiples utilisateurs l'accès à des ressources matérielles et logicielles via Internet, de manière distribuée et sous forme de services. Comme le coût pour l'utilisateur est calculé en fonction de son usage une gestion fine des ressources est primordiale. Le « cloud » doit supporter l'exécution de différentes classes de services ayant des exigences spécifiques et pouvant évoluer au cours du temps. Ces exigences concernent les aspects temporels, la sécurité, la protection, la tolérance aux fautes, la mobilité, la QoS. Une question d'actualité des systèmes distribués à grande échelle concerne les possibles convergences des systèmes P2P avec le « cloud », en particulier l'utilisation de mécanismes de base, comme les DHT. Il faut faire progresser la recherche sur les moyens d'expressions nouveaux de description des services et mettre en œuvre des stratégies de courtage et d'ordonnancement avancées. Il y a besoin d'outils de description de haut-niveau et de projections automatiques des descriptions vers des architectures effectives.

### 3.4 Sécurité, Résilience et Cryptologie

L'objectif de transparence de l'informatique ubiquitaire ne peut être atteint que si la technologie est suffisamment sûre pour que les utilisateurs puissent s'en affranchir. La résilience, la sécurité, le respect

de la vie privée et la sûreté des systèmes sont ici des problématiques transverses d'autant plus critiques que les interactions entre systèmes sont diversifiées et fréquentes. Les outils mis en œuvre proviennent de l'évaluation analytique ou expérimentale, de l'architecture matérielle et logicielle, des protocoles ou encore de l'algorithmique.

Dans les approches architecturales et protocolaires, on distinguera notamment les preuves « zero-knowledge », les protocoles d'établissement et de distribution de clés et autres mécanismes d'appairage, les protocoles de vérification de proximité/de localisation, les mécanismes d'établissement de la confiance (à base ou sans infrastructure de clés publiques), les mécanismes d'évaluation de la coopération ou du comportement. Le respect de la vie privée, qui peut concerner l'anonymat, la localisation (geo-privacy), ou encore des données personnelles, peut lui aussi être abordé par des approches algorithmiques ou cryptographiques, architecturales ou encore touchant aux données elles mêmes.

La cryptologie est un domaine vertical, à la fois proche de la complexité abstraite, de la conception d'algorithmes, de l'implémentation, jusqu'à la carte à puce. La cryptographie traditionnelle n'est pas suffisante pour répondre aux nouveaux défis des échanges sur internet et aux nouveaux scénarios tels que le vote électronique, les enchères en ligne, etc. L'industrialisation d'applications sensibles a déplacé les efforts vers des protocoles où les failles de sécurité proviennent plus souvent de la mise en œuvre des protocoles que de la solidité de la cryptographie sous-jacente. Les approches heuristiques sont remplacées par une « sécurité prouvée » et les recherches en cryptologie s'attachent à valider les schémas cryptographiques en produisant des preuves de sécurité. Enfin, la sécurité de tout protocole cryptographique reposant sur la difficulté d'un problème algorithmique, la recherche de nouveaux problèmes difficiles constitue toujours un déficit important.

## 4. Robotique

La robotique vise à doter des systèmes artificiels de capacités sensorielles, motrices et cognitives. Ces systèmes prolongent l'action de l'humain, ou le remplacent, dans des situations qui peuvent exiger plus ou moins d'autonomie. Un robot perçoit, agit, décide, apprend dans un environnement dynamique complexe, qui peut être partagé avec des êtres humains et/ou d'autres robots. Il doit alors être capable de gérer des interactions physiques et cognitives, ce qui nécessite notamment de détecter et comprendre son vis-à-vis, voire de partager l'autorité de décision ou de mener des actions conjointes. La robotique peut être présentée selon quatre thématiques scientifiques contenant chacune des verrous scientifiques et technologiques.

### 4.1 Perception, action et mouvement

Les capteurs et les actionneurs sont les composants actifs de base du robot. Leur intégration, qui fait appel à la physique des transducteurs, au traitement du signal et de l'image, à l'automatique et à l'algorithmique permet de doter le système de fonctions sensori-motrices. La recherche de nouvelles modalités perceptuelles parfois bio-inspirées, s'accompagne de problématiques de fusion de données multi-sensorielles, de modélisation de l'environnement et d'objets. L'intégration de la perception à la commande pose des problèmes variés selon que l'humain est ou n'est pas dans la boucle et selon le type de robot (terrestre à roues ou à pattes, aérien, sous-marin, manipulateur, etc.).

### 4.2 Décision, apprentissage

Ces notions désignent les capacités cognitives dont doit être doté un robot pour être autonome. Le robot décide de ses actions sur la base de l'interprétation sémantique de son état et de la variabilité de son environnement. Il raisonne sur ses actions, ce qui nécessite des fonctionnalités avancées de planification, de supervision et contrôle d'exécution en temps réel, ainsi que

d'apprentissage supervisé ou automatique. L'intégration de ces fonctionnalités requiert des architectures de contrôle et de décision complexes qui doivent être reconfigurables dynamiquement en satisfaisant de nombreuses contraintes fonctionnelles et opérationnelles.

### 4.3 Interaction et coopération

L'interaction homme-robot joue un rôle central dans la robotique de service ou dans le contexte de l'intelligence ambiante. Les problématiques portent sur la détection, l'interprétation, la fusion et la prédiction de signaux de communication (verbale et non verbale, intention, émotion), ainsi que sur la synthèse de ces signaux pour les robots ; sur la coopération et la coordination entre l'humain et le robot; sur les interfaces homme-robot (nouvelles interfaces cerveau-machine, haptiques). Les travaux sont menés avec des spécialistes des neurosciences et des sciences cognitives, et peuvent être rapprochés de ceux de la communauté de la réalité virtuelle.

### 4.4 Conception des systèmes robotiques

Les méthodes de conception doivent examiner simultanément les aspects structure et commande. Elles peuvent s'inspirer des propriétés du vivant et/ou exploiter des systèmes à base de connaissances ou les techniques avancées d'optimisation pour traiter les topologies particulières de robots à forte complexité structurelle. Elles doivent reposer sur des moyens de simulation physique hautement réalistes permettant de rendre compte d'interactions complexes (tissus mous, impacts, etc.) Il s'agit enfin de développer les constituants technologiques des systèmes robotiques du futur pour qu'ils répondent à des exigences de miniaturisation, de puissance massive, de consommation, d'intégration et de coût (sans oublier l'échelle du micromonde pour lequel de nouveaux paradigmes devront être imaginés).

## 5. Sciences et Technologies de la Décision

Ces dernières années nous assistons à l'émergence de nouvelles problématiques autour du concept de la décision: comme résultat d'interactions collectives, sociales, entre individus et/ou organisations; comme résultat de processus de décision individuels; comme résultat d'une élaboration produite par une « machine » dotée d'autonomie décisionnelle. Cette émergence est due à la très grande capacité de calcul aujourd'hui disponible et aux avancées et au rapprochement entre plusieurs axes de recherche comme la Recherche Opérationnelle, l'Intelligence Artificielle, le Choix Social ou la Théorie du Mesurage.

### 5.1 Les défis scientifiques

Les défis scientifiques se situent à l'articulation des processus de décision et d'aide à la décision et dans la manipulation des données et de l'information. Il s'agit des méthodes, des modèles, des algorithmes et des langages pour:

- apprendre (des préférences, des désirs, des contraintes, des similarités) de la part des décideurs et des parties prenantes;
- extraire des connaissances à partir de bases des données (possiblement de très grande taille) et de sources d'informations différentes (contradictoires et/ou de fiabilité variable);
- construire des synthèses d'informations (index, indicateurs, représentations visuelles) à la fois significatives et utiles;
- prendre en compte la présence d'informations incertaines, ambiguës, ainsi que des valeurs et des préférences contradictoires et garantir la robustesse à la fois des résultats et des méthodes vis-à-vis des scénarios, des incertitudes structurelles et des jeux de simulation;
- représenter les connaissances d'origines variées et issues d'interactions différentes (expertises, connaissances du terrain, opinion publique, connaissances méthodologiques) et restituer des règles et des recommandations;

- permettre la révision et la mise à jour des informations et des valeurs utilisées ainsi que la construction d'explications et des justifications ;

- résoudre des problèmes d'optimisation et/ou de satisfaction de contraintes de grande taille, de structure combinatoire, de nature stochastique et/ou dynamique et garantir la possibilité d'obtenir des « bons résultats » dans des délais raisonnables et, plus généralement, en présence de moyens de calcul définis et limités;

- restituer à l'utilisateur final des recommandations compréhensibles, légitimes, explicables, justifiables et satisfaisantes.

## 5.2 Les outils

Ce développement fait appel à des outils informatiques d'origines diverses comme:

- la programmation mathématique, (programmation en nombres entiers, optimisation combinatoire, optimisation multi-objective, programmation stochastique, dynamique, etc.);

- la théorie des graphes, la programmation par contraintes, les méta-heuristiques;

- la complexité algorithmique et la garantie des performances des algorithmes;

- la modélisation des préférences et leur agrégation en présence de critères multiples;

- les langages logiques pour la modélisation de l'hésitation, la représentation compacte de l'information, la révision ou la mise à jour ou encore l'argumentation;

- la théorie du mesurage;

- la théorie du choix social;

- les techniques de classification et d'apprentissage (numériques ou symboliques, apprentissage statistique, apprentissage logique, vector support machines, fonctions booléennes) ;

- les techniques d'extraction et de représentation des connaissances ainsi que les modèles de raisonnement;

- les méthodes de structuration, formulation et re-formulation des problèmes de décision;

- les nouveaux paradigmes de génie logiciel (systèmes multi-agents, programmation par agents), les plateformes « open-source », l'informatique distribuée, any-time, real-time.

## 5.3 Les applications

Le domaine d'application des Sciences et Technologies de la Décision est très vaste et en pleine expansion. Nous mentionnons ici quelques exemples comme la production, la logistique (commerciale, humanitaire et militaire), le développement durable, la conception, la mise en œuvre et l'évaluation des politiques publiques, la gestion de la sécurité, les réseaux de distribution (eau, énergie, communications), les services de la santé, les systèmes de recommandation et le commerce électronique, sans oublier l'industrie aéronautique, les transports, la gestion des ressources naturelles, ou encore le web-ranking, la bibliométrie, l'archéologie, la fabrication des puces électroniques et la robotique. Une telle variété d'applications positionne les Sciences et Technologies de la Décision au carrefour des différents champs disciplinaires et nécessite la construction d'un cadre méthodologique nouveau de portée plus générale, ce qui reste le défi scientifique majeur des années à venir.

## 6. Interactions

### 6.1 Image et vision

Au carrefour de plusieurs disciplines (physique, traitement du signal, intelligence artificielle et cognition, mathématiques continues et discrètes, etc.), le domaine du traitement et de l'interprétation d'images et de la vision par ordinateur se positionne au cœur des STIC. Dans l'optique d'aller de la source à l'interprétation, les évolutions portent sur les capteurs et leur usage, et sur les méthodes et les modèles. Les capteurs évoluent dans deux directions opposées, soit vers une augmentation de la résolution et de la qualité, ou vers un élargissement de bandes spectrales, comme c'est le cas en imagerie aérienne et satellitaire ou en imagerie médicale, soit au contraire vers des systèmes adaptés à de petits écrans, pour des applications multimédia sur des terminaux de capacités limitées. Les évolutions marquantes s'appuient sur des outils mathématiques, soit dans le domaine continu, soit directement dans le domaine discret.



Les interactions avec l'intelligence artificielle se développent dans diverses directions. Dans le domaine du raisonnement spatial, il s'agit de représenter les entités spatiales visibles dans les images ainsi que leurs relations spatiales, et de raisonner sur ces structures. Les approches symboliques de représentation des connaissances et de raisonnement (ontologies, systèmes à base de connaissances, logiques) sont maintenant associées aux informations, de nature plus numérique, extraites des images pour réaliser des tâches de haut niveau d'interprétation et de description du contenu des images. Les interactions avec l'intelligence artificielle sont également très fortes dans le domaine de l'apprentissage et des théories de l'incertain. Une tendance est d'introduire plus de modélisation, pour résoudre des problèmes de plus en plus complexes, en intégrant plus de connaissances. Alors que la modélisation était vue essentiellement sous l'angle de la modélisation paramétrique sur la forme ou les probabilités des signaux, les limites de ces approches pour la vision ont nécessité le développement d'autres types de modèles, s'appuyant par exemple sur la géométrie stochastique, la modélisation morphologique et/ou structurelle (graphes, grammaires, relations spatiales...). Inversement, l'analyse et l'interprétation des images conduisent à la construction de modèles (par exemple des modèles anatomiques du corps humain à partir d'images médicales).

Un domaine connexe au traitement des images est celui des informations multimédia. Ce domaine garde une composante forte sur les aspects de compression et de transmission, mais évolue également vers l'interprétation. Pour ce qui concerne le bas niveau, les paradigmes importants sont liés aux représentations parcimonieuses. Une autre évolution importante repose sur l'exploitation de la variabilité et la variété des acquisitions pour dépasser les limites individuelles des capteurs. Le domaine du codage bénéficie des techniques de codage de canal, et du rôle croissant de la théorie de l'information et

de la complexité, y compris sous leurs aspects algébriques, pour conduire à des techniques d'optimisation sans paramétrage ni critère. À plus haut niveau, l'interprétation de vidéos et de séquences d'images prend de plus en plus d'ampleur, pour le suivi ou la reconnaissance d'objets, de postures, etc. Mentionnons aussi le domaine de la synthèse d'images, de l'informatique graphique, de la géométrie numérique, de la réalité virtuelle et augmentée, en lien avec les interfaces homme-machine. Les évolutions portent sur les environnements 3D virtuels, combinant signal audio, vision, interfaces, perception, calculs sur processeur graphique (GPU), infographie, etc. Les avatars, les agents conversationnels, avec la prise en compte de plus en plus marquée des aspects émotionnels, en sont des exemples, qui sont amenés à se développer avec l'augmentation des applications de la communication homme – machine, en particulier de la communication non verbale par le geste, les expressions du visage ou les émotions.

Enfin, mentionnons quelques aspects communs aux domaines du signal, de l'image, et de la vision par ordinateur et d'autres parties de l'informatique. Nous avons déjà évoqué les interactions entre signal de parole et langage naturel, ou entre image et Sciences et Technologies de la Décision. L'apprentissage en est un exemple important, avec les applications en fouille de données et indexation. Le traitement de grosses masses de données, qui peuvent être hétérogènes, bénéficie de ces approches, mais aussi d'interactions avec le domaine des bases de données, ou encore des architectures parallèles, de l'algorithmique, des grilles de calculs, et du calcul sur GPU. Citons aussi des liens avec l'automatique et la robotique (en vision robotique, en robotique guidée par l'image, en commande référencée vision), et le domaine des interfaces homme-machine qui évolue pour introduire plus de techniques de vision et les combiner avec d'autres modalités (applications dans le domaine des interfaces homme-cerveau par exemple).

## 6.2. Représenter et apprendre les connaissances

Les connaissances sont (re-)devenues des objets d'étude fondamentaux. L'omniprésence du web sémantique et des ontologies est devenue un vrai moteur pour la recherche. Du côté théorique, il y a des recherches sur les langages de représentation des connaissances (logiques, OWL) et la formalisation du raisonnement (classification, subsomption, satisfiabilité). Les langages de représentation s'appuient essentiellement sur logiques de descriptions avec des variations qui permettent de prendre en compte des phénomènes naturels plus ou moins subtils ; des langages de représentation à base de graphes commencent à faire leur apparition. Les tendances actuelles et futures sont liées à la nécessité de construire et de disposer d'ontologies qui sont devenues des objets d'études à part entière avec des opérations propres comme l'alignement, la mise en réseau et le raisonnement distribué, la conception par composition, modulaire, par réutilisation, etc. En parallèle, d'autres problématiques montrent des liens tangibles avec le web sémantique tout en ayant des développements propres : services web, wikis et travail collaboratif, wikis sémantiques, qui ont donné parmi les plus grandes ressources de connaissances existantes (Wikipedia et autres réseaux sociaux).

La conception d'ontologies est au centre des préoccupations actuelles et futures de nos disciplines, en liaison directe avec les méthodes d'apprentissage et de fouille de données. Ainsi, les méthodes de fouille de textes, d'annotation et de recherche sémantique, vont prendre de l'importance et être capable d'appréhender les très grands volumes de documents disponibles. Ce développement donne aussi une impulsion considérable aux méthodes d'apprentissage elles-mêmes, où les préoccupations premières sont le passage à l'échelle, l'hybridation, l'applicabilité et l'intégration naturelle dans

tout système à base de connaissances. Connaissances, raisonnement et apprentissage se rejoignent. L'hybridation des méthodes d'apprentissage symboliques et numériques va connaître des développements originaux après l'effervescence des dernières années sur les méthodes numériques. Enfin, la notion de systèmes à base de connaissances ne prendra tout son sens que si ces systèmes sont capables d'apprendre comme ils commencent à être capables de raisonner.

## 6.3 Traitement des Langues Écrites, Parlées et Signées (TEPS)

En préambule, il faut souligner que cette vaste thématique de recherche, par essence pluridisciplinaire, possède de nombreux liens avec d'autres thèmes de la section 07 mais aussi avec d'autres sections du comité national (09, 27, 34). Le spectre des disciplines concernées est large et transversal : il s'étend de l'acoustique, du traitement du signal et de la modélisation statistique jusqu'à l'IA, l'algorithmique, l'apprentissage, le traitement automatique des langues, en passant par la perception, la cognition, la psychoacoustique, la linguistique, les neurosciences, etc. En outre, dans le domaine des TEPS, les méthodologies de traitement de corpus occupent une place de première importance.

Le TEPS repose sur la combinaison de plusieurs problématiques scientifiques, ce qui nécessite de construire des modèles symboliques et/ou stochastiques rendant compte des différentes facettes des systèmes de communication linguistique et de leur interrelations, qu'il s'agisse de modèles acoustico-phonétiques, de représentations lexicales, de modèles syntaxiques, de grammaires, d'ontologies, mais aussi de modèles gestuels, perceptifs et d'acquisition du langage.

De nombreux facteurs de variabilité viennent interférer avec les processus linguistiques (variabilité individuelle, environnementale,

situationnelle, sociale, liée au média...). Il s'y ajoute l'évolution naturelle des langues et des conventions d'écritures (par ex. les SMS). De tous ces facteurs, il résulte une grande complexité de la modélisation des niveaux de représentation et de leurs dépendances, et le domaine de l'algorithmique est ici crucial pour répondre aux problèmes d'apprentissage (observations incomplètes, exploitation de connaissances a priori, apprentissage discriminant) et d'optimisation des tâches de décodage (reconnaissance, transcription) et de transduction (traduction, compréhension).

A l'origine des progrès majeurs accomplis en reconnaissance de la parole dans les dernières décennies, le recours aux méthodes statistiques s'est largement répandu dans l'ensemble du domaine du TEPS, y compris pour répondre à des problèmes aussi complexes que la traduction. Parmi les sujets de recherche d'actualité, on peut citer les méthodes d'apprentissage peu ou non supervisées, l'adaptation automatique des modèles au contexte d'utilisation, la modélisation et le traitement des langues peu dotées, la caractérisation du locuteur et de ses émotions, la modélisation multi-modale, l'ensemble visant notamment à améliorer l'exploitabilité des contenus (indexation, ré-éditorialisation, distribution intelligente) et la qualité d'interaction (personnalisation, communication parlée augmentée, agents conversationnels animés, ...).

Du point de vue pratique, une part importante des débouchés se rencontre dans les domaines où les usagers d'une ou plusieurs langues sont confrontés à des difficultés de communication lors d'interactions avec des machines ou avec d'autres humains, que les TEPS peuvent alors aider à résoudre ou du moins à *médier*. La gamme des applications est très large : transcription écrite de parole, traduction automatique, localisation de logiciels, correction orthographique, recherche

d'information, résumé automatique, systèmes conversationnels, aide aux handicapés, synthèse de parole et de gestes, fouille et organisation de données sonores, etc...

Plus généralement, le TEPS s'inscrit au cœur des activités humaines de masse, qu'elles soient économiques, politiques, culturelles, professionnelles ou personnelles. Ainsi, par exemple, la traduction automatique fournit des éléments de réponse aux défis posés par le multilinguisme (notamment à l'échelle de l'Union Européenne) l'intensification du web, du e-commerce, des réseaux sociaux et la mondialisation des échanges. Dans ce contexte, l'accès aux connaissances présentes dans des contenus audio, textuels ou multimédia dépend étroitement (et stratégiquement) de la qualité des technologies de la langue écrite, parlée et signée. D'autres domaines de première importance, comme la veille en intelligence économique et la préservation du patrimoine culturel dans sa diversité linguistique sont également concernés. En résumé, les besoins technologiques nouveaux pour transcrire, analyser, comprendre, structurer, traduire, sont loin d'être satisfaits. Le domaine des TEPS apparaît donc comme porteur de développements industriels importants en réponse à une forte demande sociétale.

#### **6.4. Conception et test de circuits et systèmes intégrés.**

A l'heure des nanotechnologies, le degré de miniaturisation a atteint des densités d'intégration vertigineuses avec des circuits pouvant contenir des milliards de transistors. A ce stade, le circuit devient « système ». Cette course à l'intégration sur silicium ne s'arrête pas aux composants électroniques, de par le degré de complexité atteint. L'approche système devient inéluctable, les problématiques à caractère algorithmique, informatique ou mathématique sont au cœur

des compétences à déployer. L'augmentation des performances est devenue une logique d'évolution naturelle. Avec l'accroissement de la complexité des systèmes les problématiques de spécification, modélisation, réutilisation, optimisation de performances, tolérance aux défauts, vérification, test après fabrication, sont autant de défis à relever pour accompagner cette course à l'intégration. Les défis liés à la conception de ces systèmes relèvent pour une grande part de la section 8, mais une partie tout aussi importante relève de la section 7. Et le maître mot ici est « système complexe ». Concevoir un tel système suppose de définir des méthodologies nouvelles pour mener à bien la conception à tous les niveaux. La conception de ces systèmes complexes est par ailleurs soumise à des contraintes industrielles et économiques, on parle alors de productivité de la conception, notion qui est liée au degré de réutilisation possible à travers des bibliothèques de fonctions pouvant aller de la fonction la plus élémentaire jusqu'aux blocs virtuels les plus récents qui correspondent à un microprocesseur complet avec ses interfaces.

L'architecture matérielle de ces systèmes a beaucoup évolué. On conçoit aujourd'hui des circuits 'multi-cœurs' composés de 2, 4, 16 cœurs de processeurs sur un même substrat de silicium et il est envisagé de passer à l'échelle du 'manycore' : un système intégré composé de plus de 1000 cœurs de processeurs. Les communications intra-circuit ont dépassé le concept de fil à fil pour des 'Network-On-Chip' dans lesquels les blocs intégrés communiquent dans la puce à la manière des réseaux. C'est en fait la partie de circuit dédiée aux mémoires qui augmente le plus rapidement en proportion des autres fonctions intégrées, révolutionnant ainsi l'interface logiciel-matériel. Cette coexistence du logiciel et du matériel est telle que pratiquement toute macro-fonction complexe peut être intégrée sous forme de matériel ou sous forme de logiciel. Il en résulte que toute conception de circuit est confrontée au problème de la « co-

conception » : du partage optimal des fonctions entre logiciel et matériel. Ce constat démontre qu'il est nécessaire de poursuivre des études de manière à fournir aux concepteurs des méthodes et des modèles efficaces permettant de créer de manière sûre des systèmes hautement complexes. On distingue quatre facteurs fondamentaux qui sont à la base des recherches à poursuivre dans ce domaine :

- L'augmentation de la complexité. Les technologies CMOS utilisées pour bâtir les systèmes intégrés évoluent vers des systèmes intégrant toujours plus de composants.
- L'augmentation des performances. C'est une évolution naturelle des objets technologiques de progresser vers toujours plus de fonctionnalité, plus de vitesse, plus d'autonomie, plus de mobilité, plus de flexibilité, plus de sécurité.
- L'augmentation de l'hétérogénéité. L'intégration est un objectif constant car elle favorise l'augmentation de la vitesse et la diminution de la consommation, mais aussi l'amélioration de la fiabilité. C'est maintenant le tour des capteurs sous forme de MEMS qui cohabitent avec l'électronique de traitement sur un substrat unique de silicium.
- La fin de l'hypothèse fondatrice. La compétence du concepteur pouvait se définir jusqu'à aujourd'hui comme l'art de créer des fonctions systèmes hautement complexes par interconnexion de composants supposés parfaitement fonctionnels. Aujourd'hui, la course à la réduction des dimensions atteint ses limites et les technologies poussées dans leurs retranchements proposent de continuer à intégrer toujours plus de transistors mais en acceptant des transistors non garantis.

## 6.5 Systèmes Embarqués

Un système embarqué est un système qui interagit avec son environnement de manière autonome et qui doit être conçu et réalisé avec de très fortes contraintes : celles liées à la criticité et la taille du système, sa consommation, son prix, la qualité de services (temps de réponse,

continuité, tolérance aux pannes, etc.), la réactivité en termes de conception et d'évolution, et enfin la déployabilité sur des architectures avec des systèmes d'exploitation spécifiques et/ou multiples. La recherche et les activités de R&D dans le domaine des processeurs d'ordinateurs sont totalement dominées par les industriels américains. Les systèmes embarqués constituent un espace d'innovation où des groupes de recherche européens peuvent suivre leurs idées jusqu'à leur finalisation en R&D en collaboration avec l'industrie, comme c'est le cas avec le foisonnement de recherches autour du processeur ARM.

## 7. Interdisciplinarité

### 7.1 Systèmes d'Information

Si Internet et le web furent les déclencheurs majeurs de la globalisation de l'information par la multiplication de ses sources et de ses usages, les sciences (de la physique aux sciences de la terre en passant par la biologie sans oublier les sciences humaines) se révèlent être des domaines fortement consommateurs et producteurs d'information, qu'ils ont donc aussi besoin de traiter. Les caractéristiques transversales sont multiples : grandes masses de données et/ou de connaissances, hétérogénéité, mobilité, qualité, adaptabilité au contexte. L'information est disponible aujourd'hui à partir de nombreuses sources autonomes et hétérogènes. L'hétérogénéité de l'information se décline à tous les niveaux d'abstraction et selon divers aspects dont la combinaison est un des facteurs de complexité : hétérogénéité sémantique, de types, de formats, des représentations. L'information/les données évoluent de manière fréquente, parfois de manière permanente et intensive. La restitution de cette évolution et la qualité de l'information sont dépendantes, entre autres, du choix de granularité temporelle ainsi que de la localisation des sources. L'explosion de l'usage de capteurs produisant des données sous forme de signaux génère elle aussi de nombreux problèmes

d'acquisition et de traitement et a donné naissance aux systèmes de gestion de flots de données.

La gestion des données suit deux approches complémentaires, l'une liée aux grands volumes de données, l'autre au grand nombre de sources. Aujourd'hui la première approche est dominée par le cloud computing, des besoins en matière de réplication de données avec cohérence faible et l'apparition de systèmes de gestion de données généraux et passant à l'échelle comme BigTable. La deuxième approche utilise plutôt le paradigme P2P pour fédérer dynamiquement un grand nombre de sources, ce qui nécessite d'étendre les travaux sur les médiateurs. Ces approches sont à replacer au cœur des systèmes d'information qui ont pour objet la gestion de l'information tout au long d'une chaîne complexe : capture, saisie, stockage, production, structuration, ainsi qu'extraction, intégration, analyse, et visualisation des données. Les outils d'accès aux données doivent faire coopérer de nombreuses méthodes issues des bases de données, de la recherche d'information, de l'intelligence artificielle. Sur ce dernier aspect il convient en effet de proposer de nouveaux paradigmes qui autorisent la visualisation de données massives par l'utilisateur. Un autre aspect concerne le respect de la confidentialité des données et la sécurité des accès. Si l'on se situe moins au niveau opérationnel des systèmes d'information il est important de considérer la formalisation et la conception des systèmes dans leur dimension ubiquitaire et grande échelle. Il s'agit de proposer des méta-modèles orientés processus métier, des modèles, des langages, des outils pour appuyer des raisonnements à différents niveaux d'abstraction. Les travaux en bases de données et recherche d'information doivent intensifier l'intégration des approches pour développer les technologies permettant de rendre transparent la complexité des procédés d'analyse, d'intégration, d'extraction, d'accès aux informations, tout autant que le volume, la nature et la localisation des données accédées et ceci pour des usages allant de l'ingénierie industrielle à l'usage personnel.

## **7.2 Ethique, Déontologie, Société**

Les Sciences et technologies de l'information ont connu un développement spectaculaire, créant un nouveau monde immatériel, modifiant les relations entre les humains et leur environnement, investissant toutes les strates de la société. Des problèmes éthiques majeurs sont apparus: la cybersurveillance et la collecte de données sont susceptibles de porter atteinte au droit à la vie privée ; la diffusion d'œuvres sur internet peut contrarier le droit d'auteur ; le vote électronique institue des élections non transparentes ; l'autonomie croissante des robots soulève des questions nouvelles de responsabilité et de partage d'autorité, etc. Il est maintenant nécessaire de travailler sur l'éthique des STIC en mettant en perspectives nos sciences avec différentes disciplines des sciences humaines et sociales : philosophie, sociologie, épistémologie, mais aussi sciences juridiques et politiques. Il est nécessaire de créer des outils méthodologiques et d'élaborer des formations de haut niveau, afin que les chercheurs soient en mesure d'appréhender cette nouvelle dimension sociétale dans la pratique de leur métier, de prévoir les conséquences potentielles des recherches, et ainsi de les orienter vers de nouvelles directions. Des recherches en éthique des STIC sont déjà menées dans les pays nordiques ou anglo-saxons et ont permis l'émergence de vecteurs de diffusion des travaux de recherche. Il s'agit de prendre place dans ce nouveau champ disciplinaire trop longtemps négligé, dans lequel nous avons des compétences spécifiques en France et en particulier au CNRS.

## **7.3 Bio-informatique**

L'informatique doit fournir, avec les autres disciplines mères de la bioinformatique, les concepts nouveaux pour modéliser les processus biologiques afin de les comprendre, de les prédire, de les simuler, dans le but d'aider à répondre aux questions de la biologie

fondamentale et de ses applications. Les progrès fulgurants des techniques d'acquisition permettent d'accéder à des données biologiques de plus en plus massives et de plus en plus diverses. Par ailleurs, les bases de données biologiques contiennent une immense masse de connaissances à exploiter, mais qui sont hautement hétérogènes, changeantes, parfois peu fiables voire contradictoires. Le rapport de conjoncture de la CID 43 détaille les principaux axes de recherches fondamentales en informatique qui sont nécessaires pour répondre aux défis de l'analyse des données et de la modélisation en biologie. Nous en présentons ici une synthèse.

L'algorithmique est et restera un domaine de recherche essentiel pour la bioinformatique. Les problèmes algorithmiques posés par la biologie sont, dans l'immense majorité des cas, NP-difficiles dès lors que la modélisation tient compte de tous les paramètres « réalistes » connus. Une étude fine de ces problèmes est nécessaire pour aboutir à des approches de résolution effectives. Celles-ci peuvent être fondées notamment sur des simplifications biologiquement réalistes des modèles, sur des algorithmes d'approximation, sur des approches de complexité paramétrée, ou encore sur des heuristiques bien fondées. Les problèmes théoriquement plus faciles motivent également des recherches poussées pour passer à l'échelle (parallélisme massif, programmation sur circuits programmables (FPGA) ou cartes graphiques (GPU), structures d'indexation sophistiquées).

La classification et l'apprentissage automatiques constituent un autre domaine clé pour la bioinformatique. Ce domaine est en lien avec la statistique et la modélisation probabiliste, et avec la théorie des langages (grammaires stochastiques). Il y a pléthore de données, ce qui est un avantage ; mais aussi pléthore de paramètres, et souvent peu de données avérées expérimentalement sur lesquelles les méthodes d'apprentissage

peuvent se fonder. Le développement de méthodes pour enchaîner les étapes d'analyse à grande échelle des processus biologiques nécessite de continuer à développer des recherches dans d'autres domaines clés tels que la manipulation et l'intégration de données massives, complexes et hétérogènes, le calcul distribué et massif, l'analyse d'images et la géométrie computationnelle, la modélisation formelle de systèmes dynamiques, la visualisation. Finalement, une particularité de la recherche en informatique pour la biologie est qu'elle nécessite une réelle intrication de tous les domaines de la discipline, allant des plus théoriques, touchant les mathématiques, jusqu'aux domaines les plus proches du « hardware », en passant par la gestion des données ou l'algorithmique.

## 7.4 Développement Durable

Une profonde réflexion doit être menée sur l'impact environnemental des technologies de l'information et de la communication. Aujourd'hui, cette industrie est responsable de 2% des émissions mondiales de CO<sub>2</sub>, les seuls centres de données consomment 2 à 3% de l'électricité mondiale. Les nouveaux modes de connexion des usagers aux réseaux, permanente, mobile et à haut débit, ainsi que l'accroissement quantitatif et qualitatif des contenus échangés et stockés provoquera le doublement de l'impact énergétique de ces technologies dans la prochaine décennie. Des enjeux scientifiques profonds sont au cœur de la mutation technologique qui devient nécessaire, à savoir concevoir des systèmes matériels, protocolaires, et applicatifs optimisant efficacité énergétique et performances de service. De nombreux domaines applicatifs sont concernés, des circuits aux systèmes distribués pair à pair en passant par les terminaux fixes ou mobiles, les réseaux de communication et de capteurs sont également concernés, notamment dans les bâtiments intelligents. Cet horizon

technologique ne sera atteint qu'avec des ruptures scientifiques dans des domaines variés, et grâce à des interactions entre ces domaines : algorithmique, protocoles, optimisation combinatoire, modélisation stochastique, micro-électronique, masses de données, commande, diagnostic, etc.

D'autre part, la conception et la mise en œuvre des politiques de développement durable interpelle l'informatique sur plusieurs sujets de recherche : les méthodes d'aide à la décision et la construction d'indicateurs significatifs pour la conception, la mise en œuvre et l'évaluation des politiques publiques; le traitement d'énormes masses de données (changement climatique, qualité de l'air, des eaux et des sols), assez souvent sous forme de flot d'information continue, qui demande la mise en place d'algorithmes spécifiques de simulation et d'optimisation ainsi que l'étude de leurs performances ; le développement de méthodes d'optimisation spécifiques dans plusieurs domaines d'activité économique comme la production et distribution d'énergie, la gestion de l'eau, l'exploitation des ressources naturelles, la logistique, la planification de l'utilisation des sols.

## 8. Analyse

### 8.1 Forces

Les sciences de l'information ont une longue et forte tradition en France et constituent une discipline bien établie à la fois dans le monde académique et dans le monde industriel. La recherche menée couvre tous les aspects depuis la recherche fondamentale dans les domaines clés de l'informatique théorique, du traitement du signal et des images, de l'automatique, et de la logique, jusqu'à la recherche à l'interface avec le monde réel et en interaction avec d'autres disciplines. La communauté des chercheurs (essentiellement au sein de laboratoires universitaires) est forte et attractive pour des

chercheurs de haut niveau du monde entier (les postes ouverts par le CNRS et l'INRIA y contribuent largement). La communauté est largement structurée dans les GDR suivantes (par ailleurs plusieurs parmi ces GDR ont contribué activement à ce rapport) avec un impact très positif : GDR 673 Informatique Mathématique, GDR 717 Modélisation, Analyse et Conduite des Systèmes Dynamiques, GDR 720 Information, Signal, Images, Vision, GDR 722 Information, Interaction, Intelligence, GDR 725 Architecture, Systèmes, Réseaux, GDR 2340 Méthodes et Applications pour la Géomatique et l'Information Spatiale, GDR 2286 Mathématiques des Systèmes Perceptifs et Cognitifs, GDR 2647 STIC et Santé, GDR 2995 System On Chip – System In Package, GDR 3000 Informatique Graphique, GDR 3002 Recherche Opérationnelle, GDR 3003 Bio-informatique moléculaire, GDR 3072 Robotique, GDR 3045 Vision, GDR 3168 Génie de Programmation et du Logiciel.

## 8.2 Faiblesses

La fragmentation trop importante des laboratoires de recherche, sans stratégie associée suffisamment claire, peut être vue comme un point faible. Ce morcellement est manifeste pour les grandes institutions que sont le CNRS et l'INRIA, dont la complémentarité des missions pourrait être mieux appréhendée. Il en résulte des politiques scientifiques et de gestion de ressources insuffisamment coordonnées. Une confusion historique regrettable entre recherche en sciences de l'information et ingénierie informatique a déséquilibré notre communauté. En particuliers nos chercheurs sont surchargés de tâches qui ne devraient pas leur être dévolues, et si la ration IAT/chercheur, déjà faible au CNRS était appliqué dans les universités, notre travail serait bien plus efficace et aurait bien plus d'impact. Bien que la communauté française soit forte et structurée, elle manque encore d'ouverture européenne et internationale dans certains domaines. Les équipes Françaises sont présentes dans les projets européens mais elles pourraient l'être davantage compte tenu du

potentiel existant. Cela devient crucial au moment où plusieurs pays non Européens développent leur potentiel de recherche en sciences de l'information, et sont susceptibles de devenir, à moyen terme, plus attractifs que chaque pays pris individuellement en Europe.

## 8.3 Risques

La recherche fondamentale souffre d'un manque de moyens pérennes depuis plusieurs années. D'autre part, l'attractivité des études supérieures en sciences de l'information et des carrières de recherche stagne et pourrait même décroître à moyen terme. Cette tendance néfaste ne peut pas être compensée par le drainage de cerveaux venant d'ailleurs, puisque de plus en plus de pays atteignent ou vont atteindre un niveau d'enseignement supérieur et de recherche comparable à celui de la France et de l'Europe. La priorité « STIC » qu'affichent les gouvernements successifs depuis plusieurs années pourrait être transformée en un simple effet d'annonce si on n'impulse pas de politique nationale musclée et incitative dans le domaine : nos disciplines sont couvertes de façon importante par de nombreuses universités françaises prestigieuses, mais elles ne sont que très rarement des priorités des établissements.

## 8.4 Opportunités

La demande en sciences de l'information ne cesse de croître dans le monde entier. Les offres de postes au niveau international, à la fois dans les centres de recherche publics et privés ou directement dans l'industrie (des services) augmentent également. L'innovation industrielle dépend déjà fortement des avancées en informatique. Les technologies de l'information et de la communication vont vraisemblablement être cruciales pour contrôler le développement économique, et la société de l'information deviendra un enjeu majeur dans une perspective de globalisation dans la mesure où de plus en plus de pays y auront accès. Le développement durable, associé aux considérations éthiques, représente pour nous un enjeu considérable.