

PROSPECTIVE SCIENTIFIQUE

Document préliminaire – Synthèse des contributions des GDR

Décembre 2023

Remarques préliminaires

Ce document synthétise une analyse de prospective faite en 2022 à partir des contributions des groupements de recherche (GDR) rattachés à l'Institut des sciences informatiques et de leurs interactions du CNRS (CNRS Sciences informatiques) et des retours du conseil scientifique de ce même institut (CSI). La section « Vers une informatique plus durable » a été révisée en 2023 pour y intégrer des éléments mis en avant dans le cadre de l'année des « sciences informatiques éco-responsables » de l'institut.

Ce document a vocation à être enrichi et complété en 2024 par le rapport de prospective du conseil scientifique d'institut (CSI) et en fonction des retours des unités de recherche rattachées à l'institut qui seront consultées.

1 PRÉLIMINAIRE : ORIENTATION STRATÉGIQUE POUR LES SCIENCES INFORMATIQUES

La recherche menée au CNRS dans le cadre de l'institut des sciences informatiques et de leurs interactions est une recherche fondamentale en sciences informatiques au sens large, incluant l'algorithmique, les sciences du logiciel, du calcul et des données, l'intelligence artificielle, mais aussi le traitement du signal et des images, des langues et de la parole, l'automatique et la robotique, les systèmes embarqués ainsi que l'interaction humain-machine. Cette recherche s'inscrit dans une démarche de long terme qui vise à la fois à développer des méthodes en explorant les questions propres aux sciences informatiques et à jouer un rôle transverse en interaction avec l'ensemble des autres domaines scientifiques et en prise avec les grands enjeux socio-économiques du domaine du numérique.

Les enjeux et contributions ainsi que la durée dans lesquels l'action de l'institut s'inscrit peuvent être illustrés par la compétition lancée par le NIST (National Institute of Standards and Technology) en 2016 aux États-Unis, visant à la standardisation d'algorithmes de cryptographie post-quantique et dont les premiers résultats ont été annoncés en juillet 2022. L'enjeu à l'origine de la compétition est l'une des priorités stratégiques internationales du moment qui fait, depuis 2021, l'objet d'une stratégie d'accélération en France. Il s'agit de la sécurité informatique, déclinée ici dans une vision de long terme : se protéger contre une menace non encore avérée, mais qui pourrait se concrétiser dans les 10 à 20 prochaines années (celle de la mise au point de calculateurs quantiques susceptibles de casser les systèmes de cryptographie utilisés à l'heure actuelle). Les algorithmes acceptés pour standardisation associent, pour trois sur quatre d'entre eux, des chercheurs et enseignants-chercheurs issus d'unités mixtes de recherche (UMR) relevant de CNRS Sciences informatiques et reposent sur le principe des réseaux euclidiens. Les travaux fondateurs sur ces réseaux datent d'une vingtaine d'années et leur implémentation sous la forme de systèmes cryptographiques a fait l'objet de nombreux travaux en informatique depuis une dizaine d'années. C'est un exemple – parmi d'autres – de travaux de nature fondamentale (trouver efficacement le vecteur le plus proche d'un vecteur donné dans un sous-groupe discret d'un espace euclidien) qui ouvrent la voie à des applications à fort impact. L'annonce de la standardisation fait basculer la recherche sur ces sujets dans une nouvelle phase, prévue pour durer cinq

à six ans, dans laquelle l'implémentation sécurisée et vérifiée de ces algorithmes va devenir primordiale. C'est d'ailleurs l'enjeu principal du projet PQ-TLS (*post-quantum padlock for web browser*) du PEPR Technologies quantiques lancé en 2022. Les recherches menées au sein des laboratoires rattachés à l'institut des sciences informatiques couvrent l'ensemble des étapes de ce processus : des recherches amont sur les principes cryptographiques jusqu'à des travaux sur la cryptanalyse des implémentations logicielles ou la robustesse des implémentations matérielles. Plusieurs UMR sont également en pointe sur les questions complémentaires, mais à horizon plus lointain, liées au calcul et à l'information quantique, avec, là encore, des résultats importants obtenus pendant la période 2019-2023.

Cet exemple illustre la manière dont les recherches en sciences informatiques et les évolutions sociétales portées par les avancées des technologies numériques s'enrichissent mutuellement. D'autres enjeux socio-économiques émergents font appel à l'apprentissage automatique, au calcul à haute performance, aux interfaces humain-machine et à la robotique. Autant de domaines dont les progrès s'appuient sur des communautés scientifiques établies depuis de nombreuses années et organisées autour des laboratoires et des groupements de recherche de l'institut. L'ensemble des sciences informatiques contribue au caractère transformatif des technologies numériques.

Cette section s'appuie notamment sur une consultation effectuée auprès des 15 GDR rattachés à l'institut et du CSI. Elle prend également en compte les résultats des groupes de travail interdisciplinaires mis en place dans le cadre du contrat d'objectifs et performance¹ (COP, 2019-2023) du CNRS ainsi que les travaux de la Mission pour les initiatives transverses et interdisciplinaires (MITI) du CNRS. S'il est difficile de résumer l'ensemble de ces contributions en quelques lignes, on peut toutefois mettre en avant trois évolutions significatives.

1.1 L'apprentissage automatique s'impose de façon transverse

Non seulement les laboratoires de l'institut contribuent activement aux développements méthodologiques et à la compréhension des fondamentaux de l'apprentissage mais la quasi-totalité des disciplines scientifiques propres à l'institut se trouvent renouvelées par le développement d'approches utilisant des méthodes d'apprentissage automatique pour extraire des règles comportementales à partir de données empiriques. Selon les cas, ces méthodes servent à pallier l'absence de modèle explicite du comportement désiré (par exemple en vision par ordinateur), à compléter des modèles établis (par exemple en recherche opérationnelle) ou à extrapoler des comportements observés (par exemple en bio-informatique).

On parle couramment d'intelligence artificielle (IA) pour décrire cette évolution qui a fait l'objet de toutes les attentions durant la période 2019-2023 mais le sursaut d'activité s'est porté principalement sur le développement et l'utilisation des méthodes d'apprentissage automatique (*machine learning*), notamment d'apprentissage profond (*deep learning*) et leurs déclinaisons dites « génératives », permettant de « générer » des données extrapolées.

La même tendance s'observe dans de nombreux domaines scientifiques en dehors du champ couvert par l'institut : en physique des particules, sciences de l'univers, chimie et sciences des matériaux ou sciences de la planète et de l'environnement. Ce constat a conduit en 2021 à la création au CNRS du centre AISSAI (« IA pour la Science et Science pour l'IA »).

1.2 Montée en puissance des acteurs extra académiques

Depuis une dizaine d'années, la part prise par les grandes entreprises du secteur du numérique dans les domaines de recherche de l'institut est en forte croissance.

¹ Le bilan du COP est traité à part dans le document auto-évaluation du CNRS.

Le constat est particulièrement frappant en IA où, dans les grandes conférences du domaine (NeurIPS, ICML, etc.), la part des co-auteurs affiliés aux seuls Google et Deepmind est de l'ordre de 20% à 25% et dépasse le tiers des auteurs si l'on compte l'ensemble des entreprises du domaine. Ce que certains considèrent comme l'avancée la plus significative des années récentes (la prédiction de la structure des protéines par Alphafold puis Alphafold2) est due à Deepmind, tout comme les grands modèles de langages multilingues génériques (BERT, GPT, etc.) ont été développés par Google, Meta ou Open-AI.

Même des domaines moins exposés dans les médias comme la cryptographie n'échappent plus à cette tendance. Près de la moitié des auteurs des soumissions sélectionnés en 2022 dans la compétition NIST sur la cryptographie post-quantique apparaissent comme affiliés à des industriels plutôt qu'à des établissements académiques, même s'ils ont souvent un passé académique.

La concurrence de grands établissements universitaires, souvent extra européens, est plus classique mais elle perdure et de nouveaux acteurs apparaissent, en provenance d'Asie et du Moyen-Orient notamment.

Cette double concurrence représente un défi pour l'institut en matière d'attractivité et de rétention des talents. Elle amène également la recherche académique à redéfinir ses objectifs et son ambition en fonction de ses moyens.

1.3 Montée des attentes en matière réglementaire, éthique et environnementale

Le domaine du numérique et les changements profonds qu'il engendre dans la vie courante (omniprésence des équipements et services numériques, modifications de l'organisation du monde du travail induites par la transition numérique, bouleversements de la chronologie et de la nature des médias, des relations entre individus, etc.) soulèvent des questions éthiques sur les usages des technologies potentiellement délétères (biais, enfermement algorithmique, manipulation d'opinion, vie privée, etc.). À cela s'ajoute un questionnement croissant sur l'empreinte environnementale de ces technologies et les effets rebonds induits.

Cela se traduit notamment par des attentes croissantes en matière de réglementation (sur l'utilisation des données personnelles ou l'éthique de l'IA, notamment) et de prise en compte de l'impact environnemental du secteur du numérique. Du fait de la diversité des thématiques de recherche concernées et de la rapidité du cycle qui conduit à des innovations concrètes à partir des résultats de recherche amont (notamment lorsque les innovations reposent principalement sur des développements logiciels), ces tendances de fond amènent un nombre croissant de chercheurs à s'interroger quant aux usages de leurs recherches, tout particulièrement sous l'angle de la responsabilité éthique et de la soutenabilité environnementale. Ce mouvement est porteur de nouvelles problématiques de recherche, que l'institut accompagne, par exemple au travers du groupement de services (GDS) EcoInfo et qui sont discutées au sein du GDR Internet, IA et société dans une perspective interdisciplinaire.

2 ORIENTATIONS SCIENTIFIQUES PRIORITAIRES POUR LA PÉRIODE 2024-2028²

Les orientations scientifiques prioritaires pour la période 2024-2028 sont organisées en un ensemble de priorités thématiques, complété par une priorité transverse qui irrigue de nombreux champs scientifiques.

Au sein et au-delà de ces priorités scientifiques, CNRS Sciences informatiques reste fortement attaché à la recherche fondamentale sur ses disciplines de base (informatique fondamentale, sciences du logiciel, des données et du calcul, traitement du signal, des images, des langues, automatique, etc.) Les orientations qui suivent reposent sur un socle théorique solide qu'elles contribuent aussi à renouveler en suscitant de nouvelles

² Un résumé est donné en fin de ce document

questions fondamentales. Par ailleurs et comme illustré précédemment à propos de la cryptographie post-quantique, les ruptures technologiques de demain reposent en partie sur les avancées fondamentales d'aujourd'hui qu'il convient de soutenir dans une vision de long terme.

Sur l'ensemble de ces aspects, la problématique de la reproductibilité (des résultats, des traitements, etc.) est également destinée à prendre plus d'ampleur dans les sciences informatiques, en lien avec le mouvement général, fortement soutenu par le CNRS, vers la science ouverte dans toutes ses composantes (publications, données, logiciels).

2.1 Une orientation prioritaire transverse : apprentissage automatique et science des données

Les méthodes d'apprentissage automatique et l'approche « science des données » (accélération des découvertes scientifiques à partir de la collecte et de l'exploitation de données) représentent des enjeux transverses pour la plupart des domaines scientifiques couverts par l'institut (traitement du signal et des images, robotique, recherche opérationnelle, cybersécurité, bioinformatique, neurosciences computationnelles) et constitue un socle méthodologique pour l'ensemble des grands domaines scientifiques du CNRS dont le développement se retrouve ainsi étroitement lié aux sciences informatiques.

L'enjeu de cette priorité transverse est double.

Il s'agit premièrement de développer et d'approfondir les connaissances et outils liés à l'apprentissage automatique et à la science des données. En ce qui concerne les méthodes, on peut citer la compréhension de l'apprentissage profond, les méthodes hybrides neuro-inspirées ou informées par des modèles physiques, le traitement de données hétérogènes et de flux de données, la quantification des incertitudes, la stabilité et l'explicabilité des méthodes d'apprentissage, l'inférence de causalités, l'apprentissage distribué, la protection de la confidentialité des données ainsi que la traçabilité des données et des traitements. Au niveau des outils, un axe prioritaire concerne les très grands modèles et les architectures à base de transformeurs qui révolutionnent l'état de l'art dans de nombreux domaines – initialement surtout pour le traitement des langues mais on observe des premières avancées significatives dans d'autres domaines, notamment en vision par ordinateur. Au niveau national, le CNRS est bien positionné sur cette thématique grâce aux moyens et aux compétences réunis autour du super calculateur national Jean Zay qui ont notamment permis en 2022 l'entraînement du très grand modèle de langue Bloom dans le cadre d'une collaboration internationale public/privé coordonnée par l'entreprise Hugging Face. Il est également impliqué dans le PEPR IA.

Le second enjeu consiste à capitaliser sur ces travaux fondamentaux pour faire avancer l'ensemble des domaines applicatifs, sur les thématiques relevant des sciences informatiques et des autres instituts du CNRS. Pour ce second aspect, l'action de l'institut s'appuiera prioritairement sur les outils interdisciplinaires dédiés à l'IA mis en place au CNRS en 2021-2022, comme la commission interdisciplinaire (CID) 55 du Comité national « Sciences et données » ou le centre AISSAI.

2.2 Quatre orientations prioritaires thématiques

2.2.1 Systèmes informatiques de confiance

La place toujours croissante des outils numériques dans la société et dans le tissu économique impose d'assurer la confiance dans ces outils à plusieurs niveaux. Cela passe naturellement par la poursuite des efforts sur des thématiques socles comme la cybersécurité et la protection contre les comportements malveillants, ainsi que les sciences du logiciel pour garantir la qualité, la robustesse et la traçabilité des développements logiciels (particulièrement dans les applications critiques). La notion de confiance, abordée de manière transversale entre les sciences informatiques et les sciences humaines et sociales, est désormais présente beaucoup plus largement dans les thématiques scientifiques de l'Institut. L'explicabilité des solutions proposées par les systèmes informatiques (en particulier, ceux à base d'IA) est un enjeu capital pour en

assurer l'adoption par un large public. Il en va de même pour la garantie de propriétés d'équité (applications comme Parcoursup), de sécurité (comme pour les systèmes de vote en ligne ou de crypto-monnaies) ou d'absence de biais (de genre, ethnique, sociétal, etc.). Le traitement des données sensibles (par exemple des données de santé) soulève des défis. Il faut notamment être capable de garantir l'anonymat ou d'autres formes plus graduelles de respect du caractère personnel des données, tout en permettant un traitement de ces données dans des contextes où les données et leurs traitements peuvent être distribués sur des systèmes partiellement ouverts comme le Cloud.

2.2.2 Nouveaux paradigmes de calcul et de stockage

Le calcul et le stockage sont deux des problématiques historiques de l'informatique, impliquant aussi bien des aspects logiciels que matériels. Elles se retrouvent aujourd'hui de nouveau au cœur d'enjeux importants, souvent interdisciplinaires, qui appellent de nouvelles ruptures scientifiques.

Certaines applications (par exemple pour le climat, les sciences de l'univers ou les sciences de l'environnement) nécessitent des capacités de calcul qui sont de plusieurs ordres de grandeur supérieurs à l'état de l'art. Une réponse pourra résider dans notre capacité à exploiter des calculateurs hautes performances de nouvelle génération (Exascale), en avançant sur les questions de passage à l'échelle des codes, d'hétérogénéité des architectures, de parallélisme et de tolérance aux pannes tout en maintenant une enveloppe énergétique acceptable. Ces questions sont abordées notamment au sein du PEPR NumPEX. Une autre réponse pourra venir du développement de nouveaux paradigmes de calcul. On pense bien sûr au calcul quantique dont les résultats récents obtenus par des équipes de recherche de l'institut autour de la suprématie quantique confirment le potentiel. Les technologies quantiques (qui font l'objet depuis 2021 d'un plan de soutien national de grande ampleur) suscitent de nombreux défis qui concernent aussi bien les fondements que les aspects matériels et logiciels, avec le besoin de concevoir des langages de programmation adaptés à de nouveaux modèles de calcul.

La gestion de volumes toujours plus importants de données est également l'un des enjeux importants des prochaines années. Le traitement et la transmission des données au sein de chaînes de traitements complexes devra pouvoir se faire de façon fiable, rapide et souvent dans des contextes énergétiquement contraints (typiquement dans le cadre du *edge* ou du *fog computing*). Le stockage pourra lui aussi faire appel à des ruptures technologiques, comme amorcé récemment dans le cadre du PEPR MolecuArXiv qui a pour ambition de faire passer à l'échelle le stockage de données sur ADN et polymères artificiels, en convoquant des expertises en traitement du signal, bioinformatique et chimie.

2.2.3 Systèmes cyberphysiques en interaction avec l'humain

Parmi les thématiques de l'institut, une place particulière est accordée à celles qui ont pour objet de piloter des systèmes dans le monde physique ou d'interagir avec des sujets humains. Cela concerne notamment la robotique qu'elle soit de service, médicale ou collaborative, les interfaces humain-machine, les interactions virtuelles et à distance. Cette spécificité est due à la nécessité de prendre en compte à la fois les contraintes physiques de déplacement, de latence ou de temps de réponse des systèmes et les facteurs comportementaux humains, dans leurs composantes perceptuelles et cognitives. Ces thématiques nécessitent aussi l'accès à des plateformes expérimentales de recherche, souvent trop coûteuses ou trop complexes pour être maintenues à l'échelle d'un seul laboratoire. Elles nécessitent notamment des moyens humains importants et des compétences spécifiques.

Dans la période à venir l'action de l'institut dans ce domaine s'appuiera notamment sur les plateformes et infrastructures nationales en réseaux dont il a la charge, TIRREX (robotique) et CONTINUUM (immersion visuelle collaborative), lancées en 2021 et labellisées dans la feuille de route nationale des infrastructures de recherche en tant qu'infrastructures de recherche. Il s'appuiera aussi sur les PEPR O2R et eNSEMBLE. Au-delà de ces deux thématiques, l'institut se donne également pour objectif de renforcer les moyens et de

développer des actions de recherche fédératives sur des sujets à forts enjeux applicatifs comme le traitement automatique des langues (notamment à travers la création d'un centre de compétence pour le traitement des langues en association avec CNRS Sciences humaines et sociales) ainsi que sur les applications de la robotique et des systèmes d'interaction virtuelle dans le domaine de la santé.

2.2.4 Complexité énergétique et sobriété informatique : vers une informatique plus durable

Les dimensions environnementales et énergétiques deviennent des composantes essentielles de l'évaluation de la qualité et de la pertinence des réponses scientifiques à des problématiques sociétales complexes, et ce tout particulièrement dans le domaine du numérique. La notion d'empreinte qui englobe les aspects technologiques et sociétaux devient une donnée clé et l'enjeu est de pouvoir lui donner une consistance quantitative afin de l'estimer au travers d'une approche multicritère représentative des différents niveaux qu'elle englobe.

Là où prévalaient historiquement les seules dimensions de temps de calcul et d'espace de stockage nécessaires à la résolution d'un problème donné, la définition d'un ensemble de métriques plus complet (qualitative et quantitative) et adapté aux enjeux de sobriété devient nécessaire. En particulier, l'énergie consommée entre désormais dans l'équation, ce qui soulève de nombreuses questions de nature fondamentale et expérimentale. En premier lieu, se pose la question de la mesure de la consommation énergétique et de l'élaboration de méthodes pour quantifier les impacts environnementaux d'un système informatique, avec toute la complexité induite par les aspects matériels ou réseaux (il suffit de penser ici aux systèmes de type *edge computing*) avec des dispositifs affichant des capacités énergétiques très différentes. Il convient également, pour les systèmes cyberphysiques, de tenir compte de l'empreinte écologique de leur conception et leur fabrication, ainsi que de leur cycle de vie. Une deuxième catégorie de questions porte sur le développement de solutions qui soient les plus sobres possible tout en assurant un compromis acceptable en termes de performance et de qualité. La question des limites théoriques induites par les lois physiques permettra d'estimer les progrès atteignables en termes d'optimisation. Ces problématiques concernent aussi des contextes résilients où il faut gérer des modes dégradés ou restreints d'un point de vue énergétique, typiquement pour prendre en compte des sources d'énergie intermittentes. Enfin, la question de l'effet rebond se pose de manière globale à ces différentes problématiques : comment prévoir les changements d'usage que peuvent impliquer des nouvelles technologies moins gourmandes, plus accessibles et leur impact concret sur l'empreinte (carbone, sociale et sociétale) ?

De façon plus globale, c'est la très grande majorité des domaines des sciences informatiques (IA, informatique fondamentale, automatique, recherche opérationnelle, traitement du signal, réseaux et communication, bioinformatique, neurosciences computationnelles, adéquation algorithme architecture) qui sont appelés à renouveler leurs approches pour intégrer plus fortement ces problématiques de soutenabilité. Les questions de la frugalité des modèles d'apprentissage profond par exemple et leur capacité à être utilisés dans des contextes applicatifs à différentes échelles (du système embarqué aux bâtiments intelligents par exemple), les enjeux de portabilité, de robustesse, et de fiabilité sont différentes facettes d'un même problème scientifique qui doit répondre à des attentes fortes dans de nombreux domaines applicatifs comme la santé, les véhicules intelligents, les télécoms notamment, eux-mêmes pris dans un contexte de décarbonation.

Ce champ de recherche est encore plus vaste si l'on intègre globalement les questions de changement climatique et de résilience (optimisation des systèmes énergétique, modélisation – de l'acquisition des données à leur intégration et utilisation dans des systèmes dynamiques). À côté des enjeux au cœur des sciences informatiques, il faudra relever des défis interdisciplinaires autour de la gestion de l'énergie et des ressources dans un futur énergétique décarboné et respectueux de l'environnement. Il sera par ailleurs important d'associer pleinement les sciences humaines et sociales. En particulier le champ des sciences, techniques et société apportera un regard indispensable à une compréhension globale des usages et des impacts en termes d'empreinte des travaux scientifiques et des choix technologiques. Il offrira en outre un

point de vue nécessaire pour aborder dans toutes leurs dimensions les questions de résilience et de développement d'outils numérique low-tech prenant mieux en compte les limites planétaires, notamment en ressources.

3 RÉSUMÉ : ORIENTATIONS SCIENTIFIQUES PRIORITAIRES POUR LA PÉRIODE 2024-2028

La recherche menée dans les unités de CNRS Sciences informatiques couvre tout le spectre des sciences informatiques – l'algorithmique, les sciences du calcul, du logiciel et des données, le traitement du signal, des images, des langues et de la parole, l'automatique et la robotique mais aussi les systèmes embarqués et l'interaction humain machine – depuis leurs fondements jusqu'à leurs applications au service de la société.

Cette recherche s'inscrit dans une démarche à long terme qui vise à la fois à développer des méthodes et compétences propres aux champs des sciences informatiques et à assumer un rôle transverse, en interaction avec l'ensemble des autres domaines scientifiques et en prise avec les grands enjeux socio-économiques du domaine du numérique.

Les orientations scientifiques prioritaires pour la période 2024-2028 sont organisées en un ensemble de quatre priorités thématiques, complété par une priorité transverse qui irrigue les sciences informatiques et plus largement l'ensemble des sciences.

Au sein et au-delà de ces priorités, les recherches de nature fondamentale garantissent un socle solide pour l'existant et fertile pour de futures applications.

Quatre priorités thématiques :

- **Systèmes informatiques de confiance** : cette priorité répond au besoin d'assurer la confiance dans les outils numériques dont la place est sans cesse croissante. Elle porte sur des thématiques socles comme la cybersécurité ou les sciences du logiciel, mais va désormais au-delà afin d'apporter des réponses en matière d'explicabilité, d'équité, de traitement des données sensibles et de reproductibilité, en interaction avec les sciences humaines et sociales.
- **Calcul et stockage** : cette priorité vise à atteindre des capacités supérieures de plusieurs ordres de grandeur à l'état de l'art actuel, par exemple pour répondre aux défis des sciences de l'environnement ou de l'univers. Cela passe par l'exploitation des calculateurs hautes performances de nouvelle génération (Exascale) en avançant sur les questions de passage à l'échelle des codes, d'hétérogénéité des architectures, de parallélisme et de tolérance aux pannes, mais également par le développement de nouveaux paradigmes de calcul avec un intérêt tout particulier pour les technologies quantiques. Le traitement et le stockage de données sont également confrontés à des défis qui appellent des ruptures technologiques comme le stockage sur ADN et polymères artificiels.
- **Systèmes cyberphysiques en interaction avec l'humain** : l'objectif est de piloter des systèmes dans le monde physique ou d'interagir avec des sujets humains (notamment la robotique de service, médicale ou collaborative, les interfaces humain-machine, les interactions virtuelles et à distance). Ces recherches profiteront des plateformes et infrastructures nationales en réseaux, principalement TIRREX (robotique) et CONTINUUM (immersion visuelle collaborative) pilotées par CNRS Sciences informatiques et lancées toutes deux en 2021. L'institut se positionnera en outre comme un acteur fédérateur dans des thématiques pluridisciplinaires à travers la création d'un centre de compétence pour le traitement des langues et dans le domaine de la santé avec des applications de la robotique et des systèmes d'interaction virtuelle.

- **Complexité énergétique et sobriété informatique** : afin de prendre en compte les dimensions environnementales dans les recherches menées dans la grande majorité des sciences informatiques, il conviendra de développer les outils théoriques et des méthodes permettant de quantifier l'empreinte écologique de systèmes de plus en plus complexes et de concevoir des solutions associant logiciel et matériel, les plus sobres possible tout en assurant un bon compromis en termes de performance, robustesse et de résilience.

Une cinquième priorité de nature plus transverse porte sur **l'apprentissage automatique et la science des données** avec pour ambition de développer et d'approfondir les connaissances et outils liés à ce domaine qui est commun à la majorité des thématiques scientifiques couvertes par l'institut, et s'impose également comme un socle pour l'ensemble des sciences présentes au CNRS. Les problèmes au cœur des fondements du domaine seront attaqués tout en capitalisant sur ces travaux pour faire avancer l'ensemble des champs scientifiques et domaines applicatifs intéressés. Pour ce dernier aspect, l'institut pourra s'appuyer sur les outils interdisciplinaires qui viennent d'être créés récemment comme la CID 55 « Sciences et données » et le centre AISSAI.