

La CERNA et l'accès aux données à des fins de recherche

Guillaume Piolle
guillaume.piolle@supelec.fr

Séminaire CNIL-CAPPRIS,
19 novembre 2014

Création de la CERNA

L'informatique, et particulièrement la recherche en informatique, est confrontée à des **problématiques éthiques** nouvelles et qui lui sont propres.

Création de la CERNA

À l'automne 2012, la **CERNA** (commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique d'Allistene) est créée, à l'échelon d'Allistene, sous la présidence de Max Dauchet.

Allistene

Alliance des sciences et technologies du numérique

Structure créée en 2010 par le ministère de la recherche

Membres : CDEFI, CEA, CNRS, CPU, Inria, Institut Mines-Télécom (+ INRA, INRETS, Onera)

Objectif des alliances thématiques

- Améliorer la coordination entre les acteurs d'un même champ de recherche ;
- Bâtir une réflexion prospective sur le long terme ;
- Accroître la performance, la visibilité, le rayonnement international ;
- Participer à la valorisation de la recherche française.

Missions de la CERNA

<http://cerna-ethics-allistene.org/>

La CERNA est une commission d'éthique dédiée focalisée sur l'**éthique de la recherche** et la **recherche en éthique** dans le domaine du **numérique**.

La CERNA n'est **pas** un comité d'éthique opérationnel, qui viserait des projets spécifiques – un tel organisme relève de la responsabilité des établissements (ex. ex-COPÉ au CNRS, COERL à Inria).

La CERNA ne traite **pas** des usages.

Missions de la CERNA

La CERNA a pour mission de :

- Répondre aux questions d'ordre éthique posées par le Cocor (comité de coordination d'Allistene) ou par l'un des organismes membres, après approbation du Cocor ;
- Mener une réflexion sur l'éthique des recherches scientifiques développées en STIC ;
- Aider à exprimer les besoins spécifiques de la recherche au regard du législateur et à les couvrir dans une démarche responsable ;
- Apporter un éclairage de nature scientifique aux décideurs et à la société sur les conséquences potentielles de résultats de recherche ;
- Veiller à ce que les étudiants soient formés sur ces questions ;
- Suggérer des thèmes de recherche permettant d'approfondir la réflexion éthique dans un cadre interdisciplinaire ou de prendre en compte le résultat des réflexions éthiques.

Composition de la CERNA

		Max Dauchet Université de Lille 1	
	Bernard Alaux Cap Sciences, CCSTI Aquitaine, CNCS		Christine Balagué Institut Mines-Télécom, Conseil National du Numérique
	Danièle Bourcier CNRS, COMETS		Pierre-Antoine Chardel Institut Mines-Télécom, Université Paris Descartes
	Jean-Michel Cornu FING, Imagination for People		Gilles Dowek Inria
	Jean-Gabriel Ganascia Université P. et M. Curie, GT OnLife		Alexei Grinbaum CEA, Observatoire européen des nanotechnologies
	Claire Lobet-Maris Université de Namur		Joseph Mariani CNRS
	Alice René CNRS, CCNE		Catherine Tessier Onera
			Isabelle Bloch Institut Mines-Télécom
			Raja Chatila CNRS
			Chantal Enguehard Université de Nantes, CREIS-TERMINAL
			Claude Kirchner Inria
			Guillaume Piolle Supélec
			Sophie Vulliet-Tavernier CNIL

Fonctionnement de la CERNA

Mode de fonctionnement

- Une réunion plénière tous les deux mois ;
- Groupes de travail constitués suite à des saisines, pour formulation d'un avis consultatif :
 - *Éthique de la recherche en robotique* (Raja Chatila) ;
 - *Accès aux données à des fins de recherche en sciences et technologies du numérique* (Guillaume Piolle) ;
 - *Anonymat dans la conception de technologies numériques* (Jean-Gabriel Ganascia).

Éthique de la recherche en robotique

Premier avis de la CERNA, 26 préconisations

Thématiques des préconisations

- Autonomie et capacités décisionnelles ;
- Imitation du vivant et interaction affective et sociale ;
- Réparation et augmentation de l'humain par la machine.

Domaines d'application considérés

- Robots compagnons, domestiques, didactiques, de téléprésence ;
- Véhicules autonomes ;
- Robotique médicale ;
- Robotique de défense et de sécurité.

Accès aux données à des fins de recherche

Saisine du Cocor d'Allistene

Accès aux données à des fins de recherche en sciences et technologies du numérique

Les données ont un rôle crucial pour la recherche en sciences du numérique. Leur acquisition, leur disponibilité, leur utilisation peuvent poser des questions éthiques.

Quelles sont les questions éthiques spécifiques :

- Aux données publiquement disponibles (web social notamment) ;
- Aux données confidentielles (vidéo-surveillance, données médicales) ;
- Au stockage, au partage, à la disponibilité de ces données ?

Accès aux données à des fins de recherche

Les questions éthiques naissent à la frontière

- De nombreux projets de recherche sont **inter-** ou **trans-disciplinaires**, impliquant informatique et sociologie, psychologie, philosophie, droit, médecine. . .
- Les besoins en données proviennent souvent des exigences liées aux autres disciplines ;
- Les co-disciplines peuvent bénéficier d'une expérience antérieure quant à ces problématiques ;

Accès aux données à des fins de recherche

Les questions éthiques naissent à la frontière

- De nombreux projets de recherche sont **inter- ou trans-disciplinaires**, impliquant informatique et sociologie, psychologie, philosophie, droit, médecine. . .
- Les besoins en données proviennent souvent des exigences liées aux autres disciplines ;
- Les co-disciplines peuvent bénéficier d'une expérience antérieure quant à ces problématiques ;

- Quels enseignements tirer de ces expériences ?
- Quelles sont les problématiques spécifiques au numérique ?

Accès aux données à des fins de recherche

Les données de la recherche et la coopération internationale

L'activité des chercheurs est influencée, d'une manière ou d'une autre, par la protection des données personnelles, la loi « Informatique et Libertés », la CNIL.

Les chercheurs sont parfois amenés à s'**auto-censurer** et se trouvent souvent en décalage par rapport aux pratiques observées dans d'autres pays.

→ perception d'un certain **handicap** (en termes de liberté opérationnelle) des équipes françaises dans certains domaines.

Accès aux données à des fins de recherche

Les données de la recherche et la coopération internationale

L'activité des chercheurs est influencée, d'une manière ou d'une autre, par la protection des données personnelles, la loi « Informatique et Libertés », la CNIL.

Les chercheurs sont parfois amenés à s'**auto-censurer** et se trouvent souvent en décalage par rapport aux pratiques observées dans d'autres pays.

→ perception d'un certain **handicap** (en termes de liberté opérationnelle) des équipes françaises dans certains domaines.

- Quelle est la part réelle et la part fantasmée dans les contraintes pesant sur les chercheurs ?
- Comment mieux former et informer les chercheurs sur leur cadre réglementaire de travail ?
- Comment établir des protocoles permettant de mener sagement des recherches sur des données potentiellement problématiques ?

Accès aux données à des fins de recherche

Conservation, partage et reproductibilité

Conserver ses données expérimentales est un impératif pour permettre la validation des résultats et la mise en œuvre de recherches complémentaires.

La **reproductibilité des résultats** est, en informatique, un problème très prégnant, avec des spécificités liées à la propriété intellectuelle.

- Nécessité du **partage** des données avec d'autres chercheurs
- Impératifs de sécurité des données, principe de conservation minimale

Accès aux données à des fins de recherche

Conservation, partage et reproductibilité

Conserver ses données expérimentales est un impératif pour permettre la validation des résultats et la mise en œuvre de recherches complémentaires.

La **reproductibilité des résultats** est, en informatique, un problème très prégnant, avec des spécificités liées à la propriété intellectuelle.

- Nécessité du **partage** des données avec d'autres chercheurs
- Impératifs de sécurité des données, principe de conservation minimale

- Examen de solutions centralisées (de type CASD) ou distribuées ;
- Comment définir des critères d'accès aux données conservées ?

Accès aux données à des fins de recherche

L'utilisation des données du web social

Informations publiées sur Facebook, Yahoo, Twitter... : mises à disposition par les personnes concernées, sont-elles « rendues publiques », sont-elles librement utilisables ?

Intérêt de ces données pour de nombreux champs de recherche

Diversité des pratiques suivant les disciplines et les cultures déontologiques

Accès aux données à des fins de recherche

L'utilisation des données du web social

Informations publiées sur Facebook, Yahoo, Twitter... : mises à disposition par les personnes concernées, sont-elles « rendues publiques », sont-elles librement utilisables ?

Intérêt de ces données pour de nombreux champs de recherche

Diversité des pratiques suivant les disciplines et les cultures déontologiques

- Quelles méthodologies pour l'**information**, le recueil du **consentement** (variable en nature et en valeur), pour l'exercice des droits ?
- Comment prendre en compte le problème de la **capacité d'inférence** parfois considérable des algorithmes étudiés ?

Accès aux données à des fins de recherche

La recherche en sécurité informatique et en protection de la vie privée

Objectifs couramment poursuivis :

- Conception de meilleures techniques de **protection** ;
- Identification de **menaces** et mise à l'épreuve des moyens de protection existants.

Besoin de données spécifiques, comme des traces d'attaques informatiques ou des traces de géolocalisation

Accès aux données à des fins de recherche

La recherche en sécurité informatique et en protection de la vie privée

Objectifs couramment poursuivis :

- Conception de meilleures techniques de **protection** ;
- Identification de **menaces** et mise à l'épreuve des moyens de protection existants.

Besoin de données spécifiques, comme des traces d'attaques informatiques ou des traces de géolocalisation

Lorsqu'on étudie ou fait la démonstration d'une attaque technique contre la vie privée, devrait-on bénéficier d'un cadre juridique d'**exception** sur le traitement des données ? Est-ce nécessaire, est-ce souhaitable ?

Anonymat dans la conception de technologies numériques

Une nouvelle saisine

Question récurrente de l'**anonymisation** des jeux de données expérimentales, du travail sur des bases de données anonymes ou anonymisées

Champ de questionnement trop large pour être inclus dans l'avis « accès aux données », fait l'objet d'une nouvelle saisine de la CERNA

Anonymat dans la conception de technologies numériques

Une nouvelle saisine

Question récurrente de l'**anonymisation** des jeux de données expérimentales, du travail sur des bases de données anonymes ou anonymisées

Champ de questionnement trop large pour être inclus dans l'avis « accès aux données », fait l'objet d'une nouvelle saisine de la CERNA

- Quelle évolution historique et culturelle pour les notions d'identité, d'anonymat, de pseudonymat ?
- État de l'art techniques et possibles recommandations pour :
 - Protection de l'anonymat, gestion des identités multiples, des pseudonymes ;
 - Anonymisation, pseudonymisation, assainissement de données publiques ;
 - Mesure de la qualité d'une anonymisation et des risques de réidentification.